GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: March 11, 2003, 15:26:24 ; Search time 1159 Seconds
(without alignments)
8964.367 Million cell updates/sec

Title: US-09-936-011-1
Perfect score: 357
Sequence: 1 atgtctagggaaataaccgc.....gtaatggtgcttgtggttaa 357

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	357	100.0	357	6	AX039058	AX039058 Sequence
	2	357	100.0	4612	14	TOBNYVV2	X04197 Beet necrot
	3	355.4	99.6	1922	14	BNYVMTP1	X75575 Beet necrot
	4	347.4	97.3	357	6	AX039060	AX039060 Sequence
	5	337.8	94.6	739	6	E59890	E59890 Rhizomania-
	6	337.8	94.6	4544	14	AF197556	AF197556 Beet necr
	7	337.8	94.6	4609	6	E59889	E59889 Rhizomania-
	8	337.8	94.6	4609	14	BNYKPRO2	D84411 Beet necrot
	9	333	93.3	4544	14	AF197547	AF197547 Beet necr
	10	331.4	92.8	1922	14	BNYVMTP2	X75574 Beet necrot
	11	325	91.0	739	6	E59891	E59891 Rhizomania-
	12	217.8	61.0	4616	14	AF061869	AF061869 Beet soil
	13	68.6	19.2	3065	14	LRVBABCD	Z46351 Lychnis rin
	14	62.8	17.6	3005	14	BSBV3GNS	Z66493 Beet soil-b
	15	57.8	16.2	3591	14	PSZBETACD	M81486 Poa semilat
	16	55.2	15.5	2529	14	BVQ223598	AJ223598 Beet viru
	17	51	14.3	3215	14	BSU35769	U35769 Barley stri
	18	51	14.3	3237	14	BSU35772	U35772 Barley stri
	19	51	14.3	3289	14	HOBMSVRB	X03854 Barley stri
	20	50.8	14.2	4290	14	AF239729	AF239729 Indian pe
	21	50.6	14.2	2962	14	PXMTGBP	D30753 Potato mop-
	22	50.6	14.2	2964	14	PMO277556	AJ277556 Potato mo
	23	50.6	14.2	4504	14	PNWRNAII	L07269 Peanut clum
	24	50.2	14.1	2417	14	D86638	D86638 Broad bean
	25	49.4	13.8	3239	14	BSU35770	U35770 Barley stri
	26	49.4	13.8	3249	14	BSU35771	U35771 Barley stri
	27	43	12.0	175495	2	AC023938	AC023938 Mus muscu
c	28	43	12.0	178870	2	AC021555	AC021555 Homo sapi
c	29	43	12.0	215738	2	AC113274	AC113274 Mus muscu
	30	43	12.0	258120	2	AC114408	AC114408 Mus muscu
	31	42.8	12.0	2325	14	NVMV2	D00906 Nicotiana v
c	32	42.4	11.9	204615	2	AC112254	AC112254 Homo sapi
c	33	41	11.5	171438	2	AC099231	AC099231 Rattus no
c	34	41	11.5	172076	2	AC121895	AC121895 Mus muscu
c	35	40	11.2	32070	2	AC116983	AC116983 Dictyoste
	36	39	10.9	101149	2	AC118492	AC118492 Rattus no
	37	38.8	10.9	205793	2	AC103615	AC103615 Mus muscu
c	38	38.4	10.8	171031	2	AC104020	AC104020 Homo sapi
c	39	38.2	10.7	16241	9	AC018713	AC018713 Homo sapi
c	40	38.2	10.7	205380	2	AC037488	AC037488 Homo sapi
	41	37.8	10.6	133501	2	AC116956	AC116956 Dictyoste
c	42	37.8	10.6	179846	2	AC119387	AC119387 Rattus no
	43	37.8	10.6	194158	2	AC107761	AC107761 Mus muscu
c	44	37.4	10.5	180214	2	AC101751	AC101751 Mus muscu
	45	37.2	10.4	128433	2	AC124151	AC124151 Oryza sat

ALIGNMENTS

RESULT 1
AX039058
LOCUS AX039058 357 bp DNA linear PAT 16-NOV-2000
DEFINITION Sequence 1 from Patent EP1038961.
ACCESSION AX039058
VERSION AX039058.1 GI:11228354
KEYWORDS .
SOURCE Beet necrotic yellow vein virus.
ORGANISM Beet necrotic yellow vein virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
REFERENCE 1 (bases 1 to 357)
AUTHORS Lauber,E., Jonard,G., Guilley,H. and Richards,K.
TITLE Method for inducing viral resistance into a plant
JOURNAL Patent: EP 1038961-A 1 27-SEP-2000;
DE CENTRE NAT (FR)

```
FEATURES             Location/Qualifiers
     source          1. .357
                     /organism="Beet necrotic yellow vein virus"
                     /db_xref="taxon:31721"
     CDS             1. .357
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC16570.1"
                     /db_xref="GI:11228355"
                     /translation="MSREITARPNKNVPIVVGVCVVAFFVLLAFMQQKHKTHSGGDYG
                     VPTFSNGGIYRDGTRSADFNSNNHRAYGCGGSGGSVSSRVGQQLIVLAIVSVLIVSLL
                     QRLRSPPEHICNGACG"
BASE COUNT       89 a     58 c     93 g    117 t
ORIGIN

  Query Match           100.0%;  Score 357;  DB 6;  Length 357;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-93;
  Matches  357;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 240

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357


RESULT 2
TOBNYVV2
LOCUS       TOBNYVV2                4612 bp    RNA     linear   VRL 25-MAY-1999
DEFINITION  Beet necrotic yellow vein virus RNA-2.
ACCESSION   X04197
VERSION     X04197.1  GI:62050
KEYWORDS    coat protein; subgenome; unidentified reading frame.
SOURCE      Beet necrotic yellow vein virus.
  ORGANISM  Beet necrotic yellow vein virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
REFERENCE   1  (bases 1 to 4612)
  AUTHORS   Bouzoubaa,S., Ziegler,V., Beck,D., Guilley,H., Richards,K. and
            Jonard,G.
  TITLE     Nucleotide sequence of Beet Necrotic Yellow Vein Virus RNA-2
  JOURNAL   J. Gen. Virol. 67, 1689-1700 (1986)
COMMENT     Partial readthrough of the termination codon at position 709-711
            gives rise to 75K protein
            Data kindly reviewed (02-SEP-1987) by Bouzoubaa S.
FEATURES             Location/Qualifiers
     source          1. .4612
                     /organism="Beet necrotic yellow vein virus"
                     /strain="isolate F13"
                     /db_xref="taxon:31721"
                     /clone="pBF14, pBC2"
     repeat_unit     1. .8
                     /note="inverted repeat A"
     repeat_unit     27. .34
                     /note="inverted repeat A'"
     CDS             145. .2220
                     /note="75K protein (aa 1-691); read-through stop-codon"
                     /codon_start=1
                     /transl_except=(pos:709. .711,aa:OTHER)
                     /protein_id="CAB43511.1"
                     /db_xref="GI:4894189"
                     /translation="MSSEGRYMTWKDMSHNKFMTDRWARVSDVVSVIKQSHAMDLSKA
                     ANLSIIKTALAGLGSGWSDSNPFVSPMTRFPQTLTMYGALVLYVNLSDPEFALIMTKV
                     NTLTDSGLADNASANVRRDVVSGNKAESSGKTAGTNENSAYTLTVSLAGLAQALRLEE
                     LMWTRDKFEDRFKLPWTPVQGRTSPPGQXQLAAARVTAHIRAAKRALLYPGDSPEWVG
                     WKHFYPPPPYDVYDVPPLDIINAKLAADDIGGLVTPTPASSHGLPFEVSEEVEQANRN
                     SLWLTVGLLLAALAIGIGVTAYHRKKLQSRLRELKLLWGSTGGSGGGGGFGTELHTRA
                     TDTVSLGTTLSEHVAPAPSGLRHRPAATDSGPHEVLPFEVWVFDNLAVVYDSIGMSDL
                     FYTVREFVGVFNGEFEGLIELLESPDDDDGVYTNAPRDTAIDAYESQENYDRIDVETV
                     LIERRINLKKLLLEEAELERRERDMTMIADEEQRTLLHRLESSRVEATHVVTKAEADA
                     RAAVAMAALASKEANDYDSKMAFDRSCKEQELRLRELEVNSMPSKTERYVHTGIQGGA
                     QLAGAMAVGAMLRRGASSSSQNVSGGANIGSRSQSLTRGRSASQPLSSVGGSTRGVNN
                     NNISNTNLVRAGNSAAVSAGRSTNSGNSNFWSKLRVGEGWSKYSVERAATRAQRAIVP
                     PAPPSAPAG"
     CDS             145. .711
                     /note="coat protein; 21K (aa 1-188)"
                     /codon_start=1
                     /protein_id="CAA27791.1"
                     /db_xref="GI:62051"
                     /db_xref="SPTREMBL:Q88620"
                     /translation="MSSEGRYMTWKDMSHNKFMTDRWARVSDVVSVIKQSHAMDLSKA
                     ANLSIIKTALAGLGSGWSDSNPFVSPMTRFPQTLTMYGALVLYVNLSDPEFALIMTKV
                     NTLTDSGLADNASANVRRDVVSGNKAESSGKTAGTNENSAYTLTVSLAGLAQALRLEE
                     LMWTRDKFEDRFKLPWTPVQGRTSPPGQ"
     CDS             <712. .2220
                     /note="pot. 54K protein (aa 1-503)"
                     /codon_start=1
                     /protein_id="CAA27792.1"
                     /db_xref="GI:1334951"
                     /db_xref="SPTREMBL:Q88621"
                     /translation="QLAAARVTAHIRAAKRALLYPGDSPEWVGWKHFYPPPPYDVYDV
                     PPLDIINAKLAADDIGGLVTPTPASSHGLPFEVSEEVEQANRNSLWLTVGLLLAALAI
                     GIGVTAYHRKKLQSRLRELKLLWGSTGGSGGGGGFGTELHTRATDTVSLGTTLSEHVA
                     PAPSGLRHRPAATDSGPHEVLPFEVWVFDNLAVVYDSIGMSDLFYTVREFVGVFNGEF
                     EGLIELLESPDDDDGVYTNAPRDTAIDAYESQENYDRIDVETVLIERRINLKKLLLEE
                     AELERRERDMTMIADEEQRTLLHRLESSRVEATHVVTKAEADARAAVAMAALASKEAN
                     DYDSKMAFDRSCKEQELRLRELEVNSMPSKTERYVHTGIQGGAQLAGAMAVGAMLRRG
                     ASSSSQNVSGGANIGSRSQSLTRGRSASQPLSSVGGSTRGVNNNNISNTNLVRAGNSA
                     AVSAGRSTNSGNSNFWSKLRVGEGWSKYSVERAATRAQRAIVPPAPPSAPAG"
     CDS             2133. .3287
                     /note="42K protein (aa 1-384)"
                     /codon_start=1
                     /protein_id="CAA27793.1"
                     /db_xref="GI:62053"
                     /db_xref="SPTREMBL:Q89659"
                     /translation="MVQVQRRTGGDKGAKGNRASSAPVRSRRMTQDDWSRTHPDDIFS
                     VIEKTLVEDGYKWNGVKPGHCDWGKLKESGAIDNFRGTLEGELGKNCDLTCNAAAVKL
                     DTLQKVKTSSDWTARVGIVLGAPGVGKSTSIKNLLDKFGAKHKMVLCLPFSQLLEGVF
                     AGRLDTFLVDDLFCRSVEYGKYNTMLVDEVTRVHMCEILVLAGHLGVKNVICFGDPAQ
                     GLNYKAGSAVNYNFPIIAECYASRRFGKATADLINSSNGGGKPVVGNNEVKDSWTFEE
                     LCGKILDMSTVLVATRETQKFLLEDNIESILYSDAHGQTYDVVTIILEDEFDDAAICD
                     PNVRAVLLTRARKGGMIKMGPNIAARFKNGDFNSRGVSKSCTGDTFCEDR"
     CDS             3287. .3643
                     /note="pot. 13K protein (aa 1-117)"
                     /codon_start=1
                     /protein_id="CAA27794.1"
                     /db_xref="GI:62054"
                     /db_xref="SPTREMBL:Q88622"
                     /translation="MSREITARPNKNVPIVVGVCVVAFFVLLAFMQQKHKTHSGGDYG
                     VPTFSNGGIYRDGTRSADFNSNNHRAYGCGGSGGSVSSRVGQQLIVLAIVSVLIVSLL
                     QRLRSPPEHICNGACG"
     CDS             3627. .4025
                     /note="pot. 15K protein (aa 1-132)"
                     /codon_start=1
                     /protein_id="CAA27795.1"
                     /db_xref="GI:62055"
                     /db_xref="SPTREMBL:Q89500"
                     /translation="MVLVVKVDLSNIVLYIVAGCVVVSMLYSPFFSNDVKASSYAGAI
                     FKGSGCIMDRNSFAQFGSCDIPKHVAESITKVATKEHDVDIMVKRGEVTVRVVTLTET
```

```
                     IFIILSRLFGLAVFLFMICLMSIVWFWYHR"
     CDS             4043. .4426
                     /note="pot. 14K protein (aa 1-127)"
                     /codon_start=1
                     /protein_id="CAA27796.1"
                     /db_xref="GI:62056"
                     /db_xref="SPTREMBL:Q88623"
                     /translation="MGMVDSLCVFVGRVITEGSESVEGVERFSIKFSEWKLFTIAVFV
                     EYRELGEKECSLKDAGRLHFNVSCVKCCQKLKCKKQNKNHSKHVQNGYLRKVRNFSIL
                     GVCGDCCESFTLADEKPHVIVDPEV"
     polyA_site      4612
BASE COUNT     1200 a    760 c   1209 g   1443 t
ORIGIN

  Query Match           100.0%;  Score 357;  DB 14;  Length 4612;
  Best Local Similarity  100.0%;  Pred. No. 1.6e-93;
  Matches  357;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3287 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 3346

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3347 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 3406

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3407 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 3466

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3467 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 3526

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3527 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 3586

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3587 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 3643


RESULT 3
BNYVMTP1
LOCUS       BNYVMTP1                1922 bp    RNA     linear   VRL 01-JUN-1994
DEFINITION  Beet necrotic yellow vein mosaic virus (RGl)genomic RNA for
            transport proteins.
ACCESSION   X75575
VERSION     X75575.1  GI:496548
KEYWORDS    13K transport protein; 15K transport protein; 42K transport
            protein; trasnsport protein.
SOURCE      Beet necrotic yellow vein virus.
  ORGANISM  Beet necrotic yellow vein virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
REFERENCE   1  (bases 1 to 1922)
  AUTHORS   Koenig,R.
  JOURNAL   Unpublished
  REMARK    (sites)
REFERENCE   2  (bases 1 to 1922)
  AUTHORS   Solovyev,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-NOV-1993) Solovyev A., Institute for Biochemistry and
            Plant Virology, Messeweg 11/12, Braunschweig, Germany
FEATURES             Location/Qualifiers
     source          1. .1922
                     /organism="Beet necrotic yellow vein virus"
                     /db_xref="taxon:31721"
     CDS             1. .1155
                     /codon_start=1
                     /product="42K transport protein"
                     /protein_id="CAA53253.1"
                     /db_xref="GI:496549"
                     /db_xref="SPTREMBL:Q89659"
                     /translation="MVQVQRRTGGDKGAKGNRASSAPVRSRRMTQDDWSRTHPDDIFS
                     VIEKTLVEDGYKWNGVKPGHCDWGKLKESGAIDNFRGTLEGELGKNCDLTCNAAAVKL
                     DTLQKVKTSSDWTARVGIVLGAPGVGKSTSIKNLLDKFGAKHKMVLCLPFSQLLEGVF
                     AGRLDTFLVDDLFCRSVEYGKYNTMLVDEVTRVHMCEILVLAGHLGVKNVICFGDPAQ
                     GLNYKAGSAVNYNFPIIAECYASRRFGKATADLINSSNGGGKPVVGNNEVKDSWTFEE
                     LCGKILDMSTVLVATRETQKFLLEDNIESILYSDAHGQTYDVVTIILEDEFDDAAICD
                     PNVRAVLLTRARKGGMIKMGPNIAARFKNGDFNSRGVSKSCTGDTFCEDR"
     CDS             1155. .1511
                     /codon_start=1
                     /product="13K transport protein"
                     /protein_id="CAA53254.1"
                     /db_xref="GI:496550"
                     /db_xref="SPTREMBL:Q65682"
                     /translation="MSREITARPNKNVPIVVGVCVVAFFVLLAFMQQKHKTHSGGDYG
                     VPTFSNGGKYRDGTRSADFNSNNHRAYGCGGSGGSVSSRVGQQLIVLAIVSVLIVSLL
                     QRLRSPPEHICNGACG"
     CDS             1495. .1893
                     /codon_start=1
                     /product="15K transport protein"
                     /protein_id="CAA53255.1"
                     /db_xref="GI:496551"
                     /db_xref="SPTREMBL:Q89500"
                     /translation="MVLVVKVDLSNIVLYIVAGCVVVSMLYSPFFSNDVKASSYAGAI
                     FKGSGCIMDRNSFAQFGSCDIPKHVAESITKVATKEHDVDIMVKRGEVTVRVVTLTET
                     IFIILSRLFGLAVFLFMICLMSIVWFWYHR"
BASE COUNT      505 a    309 c    506 g    602 t
ORIGIN

  Query Match           99.6%;  Score 355.4;  DB 14;  Length 1922;
  Best Local Similarity  99.7%;  Pred. No. 4.4e-93;
  Matches  356;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1155 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 1214

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1215 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 1274

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    1275 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTAAATATAGAGACGGTACAAGATCA 1334

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1335 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 1394

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1395 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 1454

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1455 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 1511


RESULT 4
AX039060
LOCUS       AX039060                357 bp    DNA     linear   PAT 16-NOV-2000
DEFINITION  Sequence 3 from Patent EP1038961.
ACCESSION   AX039060
VERSION     AX039060.1  GI:11228356
KEYWORDS    .
SOURCE      Beet necrotic yellow vein virus.
  ORGANISM  Beet necrotic yellow vein virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
REFERENCE   1  (bases 1 to 357)
  AUTHORS   Lauber,E., Jonard,G., Guilley,H. and Richards,K.
  TITLE     Method for inducing viral resistance into a plant
  JOURNAL   Patent: EP 1038961-A 3 27-SEP-2000;
            DE CENTRE NAT (FR)
```

```
FEATURES             Location/Qualifiers
     source          1. .357
                     /organism="Beet necrotic yellow vein virus"
                     /db_xref="taxon:31721"
     CDS             1. .357
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC16571.1"
                     /db_xref="GI:11228357"
                     /translation="MSREITARPNKNVPIVVGVCVVAFFVLLAFMQQAAATHSGGDYG
                     VPTFSNGGIYRDGTRSADFNSNNHRAYGCGGSGGSVSSRVGQQLIVLAIVSVLIVSLL
                     QRLRSPPEHICNGACG"
BASE COUNT       84 a     60 c     96 g    117 t
ORIGIN

  Query Match             97.3%;  Score 347.4;  DB 6;  Length 357;
  Best Local Similarity   98.3%;  Pred. No. 8.7e-91;
  Matches  351;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120
           ||||||||||||||||||||||||||||||||||||  |  |  ||||||||||||||
Db      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGG 120

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357


RESULT 5
E59890
LOCUS       E59890                   739 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Rhizomania-resisting plant.
ACCESSION   E59890
VERSION     E59890.1  GI:18622726
KEYWORDS    JP 2000312540-A/2.
SOURCE      Beet necrotic yellow vein virus.
  ORGANISM  Beet necrotic yellow vein virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
REFERENCE   1  (bases 1 to 739)
  AUTHORS   Nomura,S., Kamitani,G., Saito,M., Kiguchi,T., Kusu,S. and Soma,C.
  TITLE     Rhizomania-resisting plant
  JOURNAL   Patent: JP 2000312540-A 2 14-NOV-2000;
            GENICHI KAMITANI,SHADANHOJIN HOKKAIDO TENSAI KYOKAI
COMMENT     OS   Beet Necrotic Yellow Vein Virus
            PN   JP 2000312540-A/2
            PD   14-NOV-2000
            PF   28-APR-1999 JP 1999122628
            PR
            PI   SHINJI NOMURA,GENICHI KAMITANI,MINAKO SAITO,TADAHIKO KIGUCHI,
            PI   SHUNZO KUSUME,
            PI   CHIHIRO SOMA
            PC   A01H5/00,C12N5/10,C12N15/09,C12N5/00,C12N15/00 CC
            FH   Key             Location/Qualifiers
            FT   source          1. .739
            FT                   /organism='Beet Necrotic Yellow Vein Virus'.
FEATURES             Location/Qualifiers
     source          1. .739
                     /organism="Beet necrotic yellow vein virus"
                     /db_xref="taxon:31721"
BASE COUNT      184 a    114 c    189 g    252 t
ORIGIN

  Query Match             94.6%;  Score 337.8;  DB 6;  Length 739;
  Best Local Similarity   96.6%;  Pred. No. 5.9e-88;
  Matches  345;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
           || ||||||||||||||||||||||||||||||||| ||| ||||||||||||| |||
Db     121 GGCGATTACGGAGTCCCAACATTTTCTAACGGTGGTAAATACAGAGACGGTACAAGGTCA 180

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 240
           ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db     181 GCTGATTTTAATAGTAATAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 240

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
           |||||||||| |||||||||| |||||||||||||| |||||||||||||| |||||||
Db     241 AGTAGTCGAGTCGGGCAGCAACTTGTTGTGTTAGCTATTGTGTCTGTGTTAATAGTATCA 300

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
           || |||||||||||||| |||||||||||||| |||||||||||||||||||||||||
Db     301 CTGTTACAACGATTAAGATCTCCACCAGAACATATTTGTAATGGTGCTTGTGGTTAA 357


RESULT 6
AF197556
LOCUS       AF197556                 4544 bp    RNA     linear   VRL 04-DEC-2000
DEFINITION  Beet necrotic yellow vein virus isolate Kas3 RNA 2 segment, partial
            sequence.
ACCESSION   AF197556
VERSION     AF197556.1  GI:11528045
KEYWORDS    .
SOURCE      Beet necrotic yellow vein virus.
  ORGANISM  Beet necrotic yellow vein virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
REFERENCE   1  (bases 1 to 4544)
  AUTHORS   Koenig,R. and Lennefors,B.L.
  TITLE     Molecular analyses of European A, B and P type sources of Beet
            necrotic yellow vein virus and detection of the rare P type in
            Kazakhstan
  JOURNAL   Arch. Virol. 145 (8), 1561-1570 (2000)
  MEDLINE   20456802
   PUBMED   11003469
REFERENCE   2  (bases 1 to 4544)
  AUTHORS   Koenig,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-1999) Institut fuer Pflanzenvirologie,
            Mikrobiologie und Biologische Sicherheit, Biologische Bundesanstalt
            fuer Land und Forstwirtschaft, Messeweg 11, D 38104 Braunschweig D
            38104, Germany
FEATURES             Location/Qualifiers
     source          1. .4544
                     /organism="Beet necrotic yellow vein virus"
                     /isolate="Kas3"
                     /db_xref="taxon:31721"
                     /country="Kazakhstan"
                     /note="closely related to P type
                     almost complete sequence of RNA 2; lacks only a few
                     nucleotides at the 5'-  and 3'-ends
                     type: P; segment: RNA 2"
     CDS             join(107. .670,674. .2179)
                     /note="contains read-through stop codon"
                     /codon_start=1
```

```
                     /product="coat protein"
                     /protein_id="AAG37092.1"
                     /db_xref="GI:11528046"
                     /translation="MSSEGRYMTWKDMSHNRFMTDRWARVSDVVSVIKQSHAMDLSKA
                     ANLSIIKTALAGLGSGWTDNNPFVSPMTRFPQTLTMYGALVLYVNLSDPEFALIMTKI
                     STLTDSGLADNASANVRRDVVSGNKAESSGKTAGTNENSAYTLTVSLAGLAQALRLEE
                     LMWTRDKFEDRLKLPWTPVQGRTSPPGQQLAAARVTAHIRAAKRALLYPGDSPEWVGW
                     KHFYPPPPYDVYDVPPLDIINAKLAADDIGGLVTPTPASSHGLPFEVSEEVEQANRNS
                     LWLTVGLLLAALAVGIGVAAYHRKKLQSRLRELKLLWGSTGGSGGGGGFDTELYMRAT
                     DTVSLGTTLSEHAASAPSGLRHRPAATDSGPHEALPFEVWAFDNLSVVYDSIGMSDLF
                     YTVREFVGVFNGEFEGLIELLESPDDDDGVYTNAPRDTAIDAYESQENYDRIDIETVL
                     LERRINLKKLLLEEAELERRERDMTMIADEEQRTLLHRLESSRVEATHAVAKAEADAR
                     AAVAMTALASKEANDYDSKMAFDRSCKEQELRLRELEVNSMPSKTERYVHTGIQGGAQ
                     LAGAMAVGAMLRRGASSSSQTVSSGANIGSRSQSVTRGRSASQPLSSVGGFTRGVNNN
                     ISNTNLVRAGNSAEVSAGRSTNSGNSNFWSKLRVGEGWSKYSVERAATRAQRAIMLPA
                     PPSAPAG"
     misc_feature    671. .673
                     /note="read-through stop codon"
     CDS             2092. .3246
                     /note="first triple gene block; necessary for virus
                     movement"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAG37093.1"
                     /db_xref="GI:11528047"
                     /translation="MVQVQRRTGGDKGAKGNYASSAPVRSRRMTQDDWSRTHPDDIFS
                     VIEKTLVEDGYKWNGVKPGHCDWGKLKESGAIDNFRGTLEGELGKNCDLTCNAAAVKL
                     DTLQKVKMSSDWTARVGIVLGAPGVGKSTSIKNLLDKFGAKHKMVLCLPFSQLLEGVF
                     AGRLDTFLVDDLFCRSVEYGKYNTMLVDEVTRVHMCEILVLAGHLGVKNVICFGDPAQ
                     GLNYKAGSAVNYNFPIIAECYASRRFGKATADLINSSNGGGKPVVGNNEVKDSWTFEE
                     LCGKILDMSTVLVATRETQKFLLEDNIESILYSDAHGQTYDVVTIILEDEFDDAAICD
                     PNVRAVLLTRARKGGMIKVGPNIAARFKNGDFNSRGVSKSCTGDTFCEDR"
     CDS             3246. .3602
                     /note="second triple gene block; necessary for virus
                     movement"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAG37094.1"
                     /db_xref="GI:11528048"
                     /translation="MSREITARPNKNVPIVVGVCVVAFFVLLAFMQQKHKTHSGGDYG
                     VPTFSNGGKYRDGTRSADFNSNNHRAYGCGGSGGSVSSRVGQQLVVLAIVSVLIVSLL
                     QRLRSPPEHICNGACG"
     CDS             3586. .3984
                     /note="third triple gene block; necessary for virus
                     movement"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAG37095.1"
                     /db_xref="GI:11528049"
                     /translation="MVLVVKVDLSNIVLYIVAGCVVVSMLYSPFFSNDVKASSYAGAV
                     FRGSGCIMDRNSFAQFGSCDIPKHVAESITKVATKEHDADVMAKRGEVTVRVVTLTET
                     IFIILSRLFGLAVFLFMICLMSIVWFWCHR"
     CDS             4002. .4385
                     /codon_start=1
                     /product="P14 protein"
                     /protein_id="AAG37096.1"
                     /db_xref="GI:11528050"
                     /translation="MGMVDSLCVFVGRVITEGSESVEGVERFSIKFSEWKLLTTAVYV
                     EYRQLGEKECSLKDVGRLHFNMSCVKCCRKLKCEKQNKNHSKHVQNGYLRKVRNFSIL
                     GVCGDCCESFTLADEKHHVIVDPEV"
BASE COUNT     1174 a    756 c   1207 g   1407 t
ORIGIN

  Query Match           94.6%;  Score 337.8;  DB 14;  Length 4544;
  Best Local Similarity  96.6%;  Pred. No. 6.5e-88;
  Matches  345;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3246 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 3305

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3306 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 3365

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
           || |||||||||||||||||||||||||||||||| |||||||||||||||||| |||
Db    3366 GGCGATTACGGAGTCCCAACATTTTCTAACGGTGGTAAATATAGAGACGGTACAAGGTCA 3425

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240
           ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    3426 GCTGATTTTAATAGTAATAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 3485

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
           |||||||||| |||||||||||| |||||||||||||| |||||||| ||||| |||
Db    3486 AGTAGTCGAGTCGGGCAGCAACTTGTTGTGTTAGCTATTGTGTCTGTGTTGATAGTATCA 3545

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
           || |||||||||||||| |||||||||||||| |||||||||||||||||||||||||
Db    3546 CTGTTACAACGATTAAGATCTCCACCAGAACATATTTGTAATGGTGCTTGTGGTTAA 3602


RESULT 7
E59889
LOCUS       E59889                  4609 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Rhizomania-resisting plant.
ACCESSION   E59889
VERSION     E59889.1  GI:18622725
KEYWORDS    JP 2000312540-A/1.
SOURCE      Beet necrotic yellow vein virus.
  ORGANISM  Beet necrotic yellow vein virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
REFERENCE   1  (bases 1 to 4609)
  AUTHORS   Nomura,S., Kamitani,G., Saito,M., Kiguchi,T., Kusu,S. and Soma,C.
  TITLE     Rhizomania-resisting plant
  JOURNAL   Patent: JP 2000312540-A 1 14-NOV-2000;
            GENICHI KAMITANI,SHADANHOJIN HOKKAIDO TENSAI KYOKAI
COMMENT     OS   Beet necrotic yellow vein virus
            PN   JP 2000312540-A/1
            PD   14-NOV-2000
            PF   28-APR-1999 JP 1999122628
            PR
            PI   SHINJI NOMURA,GENICHI KAMITANI,MINAKO SAITO,TADAHIKO KIGUCHI,
            PI   SHUNZO KUSUME,
            PI   CHIHIRO SOMA
            PC   A01H5/00,C12N5/10,C12N15/09,C12N5/00,C12N15/00 CC
            FH   Key             Location/Qualifiers
            FT   source          1. .4609
            FT                   /organism='Beet necrotic yellow vein virus'.
FEATURES             Location/Qualifiers
     source          1. .4609
                     /organism="Beet necrotic yellow vein virus"
                     /db_xref="taxon:31721"
BASE COUNT     1187 a    765 c   1221 g   1436 t
ORIGIN

  Query Match           94.6%;  Score 337.8;  DB 6;  Length 4609;
  Best Local Similarity  96.6%;  Pred. No. 6.5e-88;
  Matches  345;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3284 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 3343

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3344 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 3403

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
           || |||||||||||||||||||||||||||||||| ||| |||||||||||||| |||
Db    3404 GGCGATTACGGAGTCCCAACATTTTCTAACGGTGGTAAATACAGAGACGGTACAAGGTCA 3463

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 240
           ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db    3464 GCTGATTTTAATAGTAATAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 3523
```

```
Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
           |||||||||||| |||||||||||| |||||||||||||||| |||||||||||||| |||
Db    3524 AGTAGTCGAGTCGGGCAGCAACTTGTTGTGTTAGCTATTGTGTCTGTGTTAATAGTATCA 3583

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
           || |||||||||||||| |||||||||||||| ||||||||||||||||||||||||
Db    3584 CTGTTACAACGATTAAGATCTCCACCAGAACATATTTGTAATGGTGCTTGTGGTTAA 3640


RESULT 8
BNYKPRO2
LOCUS       BNYKPRO2                4609 bp    RNA     linear   VRL 06-FEB-1999
DEFINITION  Beet necrotic yellow vein virus genomic RNA for 42k protein, 13k
            protein, 15k protein, 14k protein, 75k protein.
ACCESSION   D84411
VERSION     D84411.1  GI:1304018
KEYWORDS    coat protein; pot. 54k protein; 75k protein; 14k protein; 15k
            protein; 13k protein; 42k protein.
SOURCE      Beet necrotic yellow vein virus (isolate:S, lab_host:Tetragonia
            expansa) cDNA to genomic RNA, clone:145-1, H20, S4, S7, pMPCRD5,
            pMPCRSX19, pMRP70, pMSD25.
  ORGANISM  Beet necrotic yellow vein virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
REFERENCE   1  (sites)
  AUTHORS   Saito,M., Kiguchi,T., Kusume,T. and Tamada,T.
  TITLE     Complete nucleotide sequence of the Japanese isolate S of beet
            necrotic yellow vein virus RNA and comparison with European
            isolates
  JOURNAL   Arch. Virol. 141 (11), 2163-2175 (1996)
  MEDLINE   97128991
REFERENCE   2  (bases 1 to 4609)
  AUTHORS   Saito,M., Kiguchi,T., Kusume,T. and Tamada,T.
  TITLE     Complete nucleotide sequence of a Japanese isolate S of beet
            necrotic yellow vein virus and comparison with European isolates
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 4609)
  AUTHORS   Kiguchi,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-1996) Tadahiko Kiguchi, Hokkaido Central
            Agricultural Experiment Station, Plant Biotechnology; Higashi 6
            Kita 15, Naganuma, Hokkaido 069-13, Japan (Tel:01238-9-2001,
            Fax:01238-9-2060)
FEATURES             Location/Qualifiers
     source          1. .4609
                     /organism="Beet necrotic yellow vein virus"
                     /isolate="S"
                     /db_xref="taxon:31721"
                     /clone="145-1, H20, S4, S7, pMPCRD5, pMPCRSX19, pMRP70,
                     pMSD25"
                     /lab_host="Tetragonia expansa"
     5'UTR           1. .144
     CDS             join(145. .708,712. .2217)
                     /note="readthrough"
                     /codon_start=1
                     /product="75k protein"
                     /protein_id="BAA18968.1"
                     /db_xref="GI:1902910"
                     /translation="MSSEGRYMTWKDMSHNKFMTDRWARVSDVVSVIKQSHAMDLSKA
                     ANLSIIKTALAGLGSGWTDNNPFVSPMTRFPQTLTMYGALVLYVNLSDPEFALIMTKV
                     STLTDSGLADNASANVRRDVVSGNKAESSGKTAGTNENSAYTLTVSLAGLAQALRLEE
                     LMWTRDKFEDRLKLPWTPVQGRTSPPGQQLAAARVTAHIRAAKRALLYPGDSPEWVGW
                     KHFYPPPPYDVYDVPPLDIINAKLAADDIGGLVTPTPASSHGLPFEVSEEVEQANRNS
                     LWLTVGLLLAALAVGIGVAAYHRKKLQSRLRELKLLWGSTGGSGGGGGFDTELYMRAT
                     DTVSLGTTLSEHAASAPSGLRHRPAATDSGPHEALPFEVWVFDNLAVVYDSIGMSDLF
                     YTVREFVGVFNGEFEGLIELLESPDDDDGVYTNAPRDTAIDAYESQENYDRIDIETVL
                     IERRINLKKLLLEEAELERRERDMTMIADEEQRTLLHRLESSRVEATHAVAKAEADAR
                     AAVAMAALASKEANDYDSKMAFDRSCKEQELRLRELEVNSMPSKTERYVHTGIQGGAQ
                     LAGAMAVGAMLRRGAGSSSQTVSSGANIGSRSQSLTRGRSASQPLSSVGGSTRGVNNN
                     ISNTNLVRAGNSAEVSAGRSTNSGNSNFWSKLRVGEGWSKYSVERAATRAQRAIVLPA
                     PPSAPAG"
     mat_peptide     145. .708
                     /product="coat protein"
     mat_peptide     712. .2214
                     /product="pot. 54k protein"
     CDS             2130. .3284
                     /codon_start=1
                     /product="42k protein"
                     /protein_id="BAA12342.1"
                     /db_xref="GI:1304022"
                     /translation="MVQVQRRTGGDKGAKGNRASSAPVRSRRMTQDDWSRTHPDDIFS
                     VIEKTLVEDGYKWNGVKPGHCDWGKLKESGAIDNFRGTLEGELGKNCDLTCNAAAVKL
                     DTLQKVKMSSDWTARVGIVLGAPGVGKSTSIKNLLDKFGAKHKMVLCLPFSQLLEGVF
                     AGRLDTFLVDDLFCRSVGYGKYNTMLVDEVTRVHMCEILVLAGHLGVKNVICFGDPAQ
                     GLNYKAGSAVNYNFPIIAECYASRRFGKATADLINSSNGGGKPVVGNNEVKDSWTFEE
                     LCGKILDMSTVLVATRETQKFLLEDNIESILYSDAHGQTYDVVTIILEDEFDDAAICD
                     PNVRAVLLTRARKGGMIKMGPNIAARFKNGDFNSRGVSKSCTGDTFCEDR"
     CDS             3284. .3640
                     /codon_start=1
                     /product="13k protein"
                     /protein_id="BAA12343.1"
                     /db_xref="GI:1304023"
                     /translation="MSREITARPNKNVPIVVGVCVVAFFVLLAFMQQKHKTHSGGDYG
                     VPTFSNGGKYRDGTRSADFNSNNHRAYGCGGSGGSVSSRVGQQLVVLAIVSVLIVSLL
                     QRLRSPPEHICNGACG"
     CDS             3624. .4022
                     /codon_start=1
                     /product="15k protein"
                     /protein_id="BAA12344.1"
                     /db_xref="GI:1304024"
                     /translation="MVLVVKVDLSNIVLYIVAGCVVVSMLYSPFFSNDVKASSYAGAV
                     FKGSGCIMDRNSFAQFGSCDIPKHVAESITKVATKEHDADIMVKRGEVTVRVVTLTET
                     LFIILSRLFGLAVFLFMICLMSIVWFWCHR"
     CDS             4034. .4423
                     /codon_start=1
                     /product="14k protein"
                     /protein_id="BAA12345.1"
                     /db_xref="GI:1304025"
                     /translation="MSMGMVDSLCVFVGRVITEGSESVEGVERFSIKFSDWKLFTTAV
                     YVEYRQLGEKECSLKDVGRLHFNMSCVKCCQKLKCKKQNKNHSKHVQNGYLRKVRNFS
                     ILGVCGDCCESFTLADEKHHVIVDPEV"
BASE COUNT     1187 a    765 c   1221 g   1436 t
ORIGIN

  Query Match           94.6%;  Score 337.8;  DB 14;  Length 4609;
  Best Local Similarity  96.6%;  Pred. No. 6.5e-88;
  Matches  345;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3284 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 3343

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3344 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 3403

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
           || |||||||||||||||||||||||||||||| ||| ||||||||||||| |||
Db    3404 GGCGATTACGGAGTCCCAACATTTTCTAACGGTGGTAAATACAGAGACGGTACAAGGTCA 3463

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 240
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    3464 GCTGATTTTAATAGTAATAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 3523

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
           |||||||||||| |||||||||||| |||||||||||||||| |||||||||||||| |||
Db    3524 AGTAGTCGAGTCGGGCAGCAACTTGTTGTGTTAGCTATTGTGTCTGTGTTAATAGTATCA 3583

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
           || |||||||||||||| |||||||||||||| ||||||||||||||||||||||||
Db    3584 CTGTTACAACGATTAAGATCTCCACCAGAACATATTTGTAATGGTGCTTGTGGTTAA 3640


RESULT 9
AF197547
LOCUS       AF197547                4544 bp    RNA     linear   VRL 04-DEC-2000
DEFINITION  Beet necrotic yellow vein virus isolate F75 RNA 2 segment, partial
```

```
            sequence.
ACCESSION   AF197547
VERSION     AF197547.1  GI:11528023
KEYWORDS    .
SOURCE      Beet necrotic yellow vein virus.
  ORGANISM  Beet necrotic yellow vein virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
REFERENCE   1  (bases 1 to 4544)
  AUTHORS   Koenig,R. and Lennefors,B.L.
  TITLE     Molecular analyses of European A, B and P type sources of Beet
            necrotic yellow vein virus and detection of the rare P type in
            Kazakhstan
  JOURNAL   Arch. Virol. 145 (8), 1561-1570 (2000)
  MEDLINE   20456802
   PUBMED   11003469
REFERENCE   2  (bases 1 to 4544)
  AUTHORS   Koenig,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-1999) Institut fuer Pflanzenvirologie,
            Mikrobiologie und Biologische Sicherheit, Biologische Bundesanstalt
            fuer Land und Forstwirtschaft, Messeweg 11, D 38104 Braunschweig D
            38104, Germany
FEATURES             Location/Qualifiers
     source          1. .4544
                     /organism="Beet necrotic yellow vein virus"
                     /isolate="F75"
                     /db_xref="taxon:31721"
                     /country="France"
                     /note="almost complete sequence of RNA 2; lacks only a few
                     nucleotides at the 5'-  and 3'-ends
                     type: P; segment: RNA 2"
     CDS             join(107. .670,674. .2179)
                     /note="contains read-through stop codon"
                     /codon_start=1
                     /product="coat protein"
                     /protein_id="AAG37079.1"
                     /db_xref="GI:11528024"
                     /translation="MSSEGRYMTWKDMSHNRFMTDRWARVSDVVSVIKQSHAMDSSKA
                     ANLSIIKTALAGLGSGWTDNNPFVSPMTRFPQTLTMYGALVLYVNLSDPEFALIMTKI
                     STLTDSGLADNASANVRRDVVSGNKAESSGKTAGTNENSAYTLTVSLAGLAQALRLEE
                     LMWTRDKFEDRLKLPWTPVQGRTSPPGQQLAAARVTAHIRAAKRALLYPGDSPEWVGW
                     KHFYPPPPYDVYDVPPLDIINAKLAADDIGGLVTPTPASSHGLPFEVSEEVEQANRNS
                     LWLTVGLLLAALAVGIGVAAYHRKKLQSRLRELKLLWGSTGGSGGGGGLDTELYMRAT
                     DTVSLGTTPSEHAALAPSGLRHRPAATDSGPHEALPFEVWVFDNLAVVYDSIGMSDLF
                     YTVREFVGVFNGEFEGLIELFESPDDDDGVYTNAPRDTAIDAYESQENYDRIDIETVL
                     LERRINLKKLLLEEAELERRERDMTMIADEEQRTLLHRLESSRVEAAHAVAKAEADAR
                     AAVAMAALASKEANDYDSKMAFDRSCKEQELRLRELEVNSMPSKTERYVHTGIQGGAQ
                     LAGAMAVGAMLRRGASSSSQTVSSGANIGSRSQSLTRGRSASQPLSSVGGSTRGVNNN
                     ISNTNLVRAGNSAEVSAGRSTNSGNSNFWSKLRVGEGWSKYSVERAATRAQRAITLPA
                     PPSAPAG"
     misc_feature    671. .673
                     /note="read-through stop codon"
     CDS             2092. .3246
                     /note="first triple gene block; necessary for virus
                     movement"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAG37080.1"
                     /db_xref="GI:11528025"
                     /translation="MVQVQRRTGGDKGAKGNYASSAPVRSRRMTQDDWSRTHPDDIFS
                     VIEKTLVEDGYKWNGVKPGHCDWGKLKESGAIDNFRGTLEGELGKNCDLPCNAAAVKL
                     DTLQKVKMSSDWTARVGIVLGAPGVGKSTSIKNLLDKFGAKHKMVLCLPFSQLLEGVF
                     AGRLDTFLVDDLFCRSVEYGKYNTMLVDEVTRVHMCEILVLAGHLGVKNVICFGDPAQ
                     GLNYKAGSAVNYNFPIIAECYASRRFGKATADLINSSNGGGKPVVGNNEVKDSWTFEE
                     LCGKILDMSTVLVATRETQKFLLEDNIESILYSDAHGQTYDVVTIILEDEFDDAAICD
                     PNVRAVLLTRARKGGMIKMGPNIAARFKNGDFNSRGVSKSCTGDTFCEDR"
     CDS             3246. .3602
                     /note="second triple gene block; necessary for virus
                     movement"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAG37081.1"
                     /db_xref="GI:11528026"
                     /translation="MSREITARPNKNVPIVVGVCVVAFFVLLAFMQQKHKTHSGGDYG
                     VPTFSNGGKYRDRTRSADFNSNNHRAYGCGGSGGSVSSRVGQQLVVLAIVSVLIVSLL
                     QRLRSPPEHICNGACG"
     CDS             3586. .3984
                     /note="third triple gene block; necessary for virus
                     movement"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAG37082.1"
                     /db_xref="GI:11528027"
                     /translation="MVLVVKVDLSNIVLYIVAGCVVVSMLYSPFFSNDVKASSYAGAV
                     FRGSGCIMDRNSFAQFGSCDIPKHVAESITKVATKEHDADVMVKRGEVTVRVVTPTET
                     IFIILSRLFGLAVFLFMICLMSIVWFWCHR"
     CDS             4002. .4385
                     /codon_start=1
                     /product="P14 protein"
                     /protein_id="AAG37083.1"
                     /db_xref="GI:11528028"
                     /translation="MGMVDSLCVFVGRVITEGSESVEGVERFSIKFSEWKLFTTAVYV
                     EYRQLGEKECSLKDVGRLHFNMSCVKCCRKLKCKKQNKNHSKHVQNGYLRKVRNFSIL
                     GVCGDCCESFTLADEKHHVIVDPEV"
BASE COUNT     1174 a    766 c   1205 g   1399 t
ORIGIN

  Query Match             93.3%;  Score 333;  DB 14;  Length 4544;
  Best Local Similarity   95.8%;  Pred. No. 1.7e-86;
  Matches  342;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    3246 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTCTGT 3305

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3306 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 3365

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
           || |||||||||||||||||||||||||||||||||| |||||||||| | ||||| |||
Db    3366 GGCGATTACGGAGTCCCAACATTTTCTAACGGTGGTAAATATAGAGACCGCACAAGGTCA 3425

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 240
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    3426 GCTGATTTTAATAGTAATAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 3485

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
           |||||||||||| |||||||||||| |||||||||||||||| |||||||| ||||| |||
Db    3486 AGTAGTCGAGTCGGGCAGCAACTTGTTGTGTTAGCTATTGTGTCTGTGTTGATAGTATCA 3545

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
           || ||||||||||||| |||||||||||||| ||||||||||||||||||||||||||
Db    3546 CTGTTACAACGATTAAGATCTCCACCAGAACATATTTGTAATGGTGCTTGTGGTTAA 3602


RESULT 10
BNYVMTP2
LOCUS       BNYVMTP2                1922 bp    RNA     linear   VRL 01-AUG-1994
DEFINITION  Beet necrotic yellow vein mosaic virus (Yu2) genomic RNA for
            transport proteins.
ACCESSION   X75574
VERSION     X75574.1  GI:496552
KEYWORDS    13K transport protein; 15K transport protein; 42K transport
            protein; transport protein.
SOURCE      Beet necrotic yellow vein virus.
  ORGANISM  Beet necrotic yellow vein virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
REFERENCE   1  (bases 1 to 1922)
  AUTHORS   Kruse,M., Koenig,R., Hoffmann,A., Kaufmann,A., Commandeur,U.,
            Solovyev,A.G., Savenkov,I. and Burgermeister,W.
  TITLE     Restriction fragment length polymorphism analysis of reverse
            transcription-PCR products reveals the existence of two major
            strain groups of beet necrotic yellow vein virus
  JOURNAL   J. Gen. Virol. 75 (Pt 8), 1835-1842 (1994)
  MEDLINE   94321971
```

```
   PUBMED   7913953
REFERENCE   2  (bases 1 to 1922)
  AUTHORS   Solovyev,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-NOV-1993) Solovyev A., Institute for Biochemistry and
            Plant Virology, Messeweg 11/12, Braunschweig, Germany
FEATURES             Location/Qualifiers
     source          1. .1922
                     /organism="Beet necrotic yellow vein virus"
                     /db_xref="taxon:31721"
     CDS             1. .1155
                     /codon_start=1
                     /product="42K transport protein"
                     /protein_id="CAA53250.1"
                     /db_xref="GI:496553"
                     /db_xref="SPTREMBL:Q65683"
                     /translation="MVQVQRRTGGDKGAKGNRASSAPVRSRRMTQDDWSRTHPDDIFS
                     VIEKTLVEDGYKWNGVKPGHCDWGKLKESGAVDNFRGTLEGELGKNCDLTCNAAAVKL
                     DTLQKVKMSSDWTARVGIVLGAPGVGKSTSIKNLLDKFGAKHKMVLCLPFSQLLEGVF
                     AGRLDTFLVDDLFCRSVEYGKYNTMLVDEVTRVHMCEILVLAGHLGVKNVICFGDPAQ
                     GLNYKAGSAVNYNFPIIAECYASRRFGKATADLINSSNGGGKPVVGNNEVKDSWTFEE
                     LCGKILDMSTVLVATRETQKFLLEDNIESILYSDAHGQTYDVVTIILEDEFDDAAICD
                     PNVRAVFLTRARKGGMIKMGPNIAARFKNGDFNSRGVSKSCTGDTFCEDR"
     CDS             1155. .1511
                     /codon_start=1
                     /product="13K transport protein"
                     /protein_id="CAA53251.1"
                     /db_xref="GI:496554"
                     /db_xref="SPTREMBL:Q65684"
                     /translation="MSREITARPNKNVPIVVGVCVVAFFVLLAFMQQKHKTHAGGDYG
                     VPTFSNGGKYRDGTRSADFNSNNHRAYGCGGSGGSVSSRVGQQLVVLAIVSVVIVLLL
                     QRLRSPPEHICNGACG"
     CDS             1495. .1893
                     /codon_start=1
                     /product="42K transport protein"
                     /protein_id="CAA53252.1"
                     /db_xref="GI:496555"
                     /db_xref="SPTREMBL:Q65685"
                     /translation="MVLVVKVDLSNIVLYIVAGCVVVSMLYSPFFSNDVKASSYAGAV
                     FKGSGCIMDRNSFAQFGSCDIPKHVAESITKVATKEHDADIMVKRGDVTVRVVTLTET
                     IFIILSRLFGLAVFLFMICLMSIVWFWCHR"
BASE COUNT      505 a    309 c    509 g    599 t
ORIGIN

  Query Match             92.8%;  Score 331.4;  DB 14;  Length 1922;
  Best Local Similarity   95.5%;  Pred. No. 4.6e-86;
  Matches  341;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    1155 ATGTCTAGGGAAATAACCGCCCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 1214

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db    1215 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATGCTGGG 1274

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
           || |||||||||| |||||||||||||||||||| |||||||||||||||||| |||
Db    1275 GGCGATTACGGAGTTCCAACATTTTCTAACGGTGGTAAATATAGAGACGGTACAAGGTCA 1334

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    1335 GCTGATTTTAATAGTAATAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 1394

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
          |||||||||| |||||||||| |||||||||||||||| |||||| ||||||| | |
Db    1395 AGTAGTCGAGTCGGGCAGCAACTTGTTGTGTTAGCTATTGTGTCTGTGGTAATAGTATTA 1454

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
           || |||||||||||||| |||||||||||||| ||||||||||||||||||||||||
Db    1455 CTGTTACAACGATTAAGATCTCCACCAGAACATATTTGTAATGGTGCTTGTGGTTAA 1511
```

```
RESULT 11
E59891
LOCUS       E59891                   739 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Rhizomania-resisting plant.
ACCESSION   E59891
VERSION     E59891.1  GI:18622727
KEYWORDS    JP 2000312540-A/3.
SOURCE      Beet necrotic yellow vein virus.
  ORGANISM  Beet necrotic yellow vein virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
REFERENCE   1  (bases 1 to 739)
  AUTHORS   Nomura,S., Kamitani,G., Saito,M., Kiguchi,T., Kusu,S. and Soma,C.
  TITLE     Rhizomania-resisting plant
  JOURNAL   Patent: JP 2000312540-A 3 14-NOV-2000;
            GENICHI KAMITANI,SHADANHOJIN HOKKAIDO TENSAI KYOKAI
COMMENT     OS   Beet Necrotic Yellow Vein Virus
            PN   JP 2000312540-A/3
            PD   14-NOV-2000
            PF   28-APR-1999 JP 1999122628
            PR
            PI   SHINJI NOMURA,GENICHI KAMITANI,MINAKO SAITO,TADAHIKO KIGUCHI,
            PI   SHUNZO KUSUME,
            PI   CHIHIRO SOMA
            PC   A01H5/00,C12N5/10,C12N15/09,C12N5/00,C12N15/00 CC
            FH   Key             Location/Qualifiers
            FT   mutation        (151). .(174).
FEATURES             Location/Qualifiers
     source          1. .739
                     /organism="Beet necrotic yellow vein virus"
                     /db_xref="taxon:31721"
BASE COUNT      182 a    117 c    187 g    253 t
ORIGIN

  Query Match             91.0%;  Score 325;  DB 6;  Length 739;
  Best Local Similarity   94.4%;  Pred. No. 3.3e-84;
  Matches  337;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
           || |||||||||||||||||||||||||||| || || ||  ||||  |   |||| |||
Db     121 GGCGATTACGGAGTCCCAACATTTTCTAACGCTGCTAAATTCAGAGCTGCAGCAAGGTCA 180

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 240
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db     181 GCTGATTTTAATAGTAATAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 240

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
          |||||||||| |||||||||| |||||||||||||||| |||||||||||||| |||
Db     241 AGTAGTCGAGTCGGGCAGCAACTTGTTGTGTTAGCTATTGTGTCTGTGTTAATAGTATCA 300

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
           || |||||||||||||| |||||||||||||| ||||||||||||||||||||||||
Db     301 CTGTTACAACGATTAAGATCTCCACCAGAACATATTTGTAATGGTGCTTGTGGTTAA 357
```

```
RESULT 12
AF061869
LOCUS       AF061869                 4616 bp    RNA     linear   VRL 26-MAR-2002
DEFINITION  Beet soil-borne mosaic virus RNA2, complete genome.
ACCESSION   AF061869
VERSION     AF061869.1  GI:3136264
KEYWORDS    .
SOURCE      beet soil-borne mosaic virus.
  ORGANISM  beet soil-borne mosaic virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
REFERENCE   1  (bases 1 to 4616)
```

```
  AUTHORS   Lee,L., Telford,E.B., Batten,J.S., Scholthof,K.B. and Rush,C.M.
  TITLE     Complete nucleotide sequence and genome organization of Beet
            soilborne mosaic virus, a proposed member of the genus Benyvirus
  JOURNAL   Arch. Virol. 146 (12), 2443-2453 (2001)
  MEDLINE   21670014
   PUBMED   11811691
REFERENCE   2  (bases 1 to 4616)
  AUTHORS   Batten,J.S., Rush,C.M. and Scholthof,K.-B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-APR-1998) Plant Pathology and Microbiology, Texas A&M
            University, 120 Peterson Building, College Station, TX 77843-2132,
            USA
FEATURES             Location/Qualifiers
     source          1. .4616
                     /organism="beet soil-borne mosaic virus"
                     /strain="EA"
                     /db_xref="taxon:76343"
                     /note="RNA2; Furovirus; submitter suggested
                     classification: genus Benevirus, family Tubiviridae"
     misc_feature    1
                     /note="5'cap"
     CDS             146. .2206
                     /note="translational readthrough of the 21 kDa coat
                     protein gene; sequence contains an amber codon at amino
                     acid (Tyr) position 192, which is a putative readthrough
                     codon (UAG)"
                     /codon_start=1
                     /transl_except=(pos:719. .721,aa:Tyr)
                     /product="75KDa protein"
                     /protein_id="AAC18569.1"
                     /db_xref="GI:3172384"
                     /translation="MVDEGRYMTWKDATHNKLMTDRWARVSSVQATINQALSLDLSKA
                     ASLPVIKANFSSLGANWSEDGDVFVAPMVRFPVTLNQFGALVLWLNLSDPAFAILMTK
                     VFTLTNAGGADNASASSRREVVAGSRPETSGKSVGTDESGSYTLTTTLQAFNQAIRLP
                     EVLWTRDKFEREWNLPWTPTRGVTPAVPRTNYQLNAARTMAHIRAALKARLYPGDSIE
                     WVGWVHSYPPPPYDGYDVPSLDIINEKLAADDVGGLVLPTPDIPGGPNFEVSEDIEQT
                     GRNGLWLMVGLLLAALAVGVGTAAYHRRKLRSRLVELKSLWFSRGSGGGGDFTVELPT
                     RTTDAFSLGTTISEHAPPPVGTMRHRSTTATDSHEALPFETWVYLNLSRVYFSIGCSD
                     LFYATREFVTVFNGEFEGLIELLEASDEDDGVYTDAESAVIVGTSAIVNDHDQFDLNT
                     VILEKRIKFKRLSLEEAELNRQESVTIEFADEERRTLMHKLESDRLEATHKVNKAAAD
                     AEAALSVAVLAAKEARTYDDKLAFDRACKEKELRLRELEVERMPGKTERYIHTGIQGG
                     AQLAGALAVGSMLRTRGVSNITPSNASGSPTGIGRVRSASQPVPTSLSGGSVLGANNN
                     NSVGSVSSGSRVESLGYNSAIRNHNTFWSKLRIGSEWRQNSISRMQARPLAGGGIPPS
                     APAV"
     CDS             146. .721
                     /note="21KDa"
                     /codon_start=1
                     /product="coat protein"
                     /protein_id="AAC18568.1"
                     /db_xref="GI:3136265"
                     /translation="MVDEGRYMTWKDATHNKLMTDRWARVSSVQATINQALSLDLSKA
                     ASLPVIKANFSSLGANWSEDGDVFVAPMVRFPVTLNQFGALVLWLNLSDPAFAILMTK
                     VFTLTNAGGADNASASSRREVVAGSRPETSGKSVGTDESGSYTLTTTLQAFNQAIRLP
                     EVLWTRDKFEREWNLPWTPTRGVTPAVPRTN"
     CDS             2125. .3273
                     /codon_start=1
                     /product="42KDa protein"
                     /protein_id="AAC18570.1"
                     /db_xref="GI:3136267"
                     /translation="MAPEQHKQNASETASGRRNSSVRSRGMSKDDWSVTHPDDVFSII
                     EKTLVEDGYKWHGVKPGHCDWDKLEQSGAIKNFKGTLEGEVDSSCSLTCNAAAIKLDI
                     VERLDVSSDWSARVGIVLGAPGVGKSTSIKHILDTYGSRYKMVLCLPVKQLLDGVFSG
                     RMDTFLIDDIFSRSVDYGKYHTMLVDEITRVHMCEVLVLAGYLGIKNVICFGDPAQGI
                     NFKAGSAVNYNFPVIAECYSSRRFGVATADLINSCNGGGKSVVGNNDVKDNWTFEELC
                     GKIEEMSTVLVATHATKEFLADDGIEAVYYEDAQGMTYDVVTIVLKDEFDDDAICDSN
                     VRAVLLTRARKGGLLKVDPNIAARFKNGDFNSRGVSKACTGDTFCEDR"
     CDS             3273. .3629
                     /codon_start=1
                     /product="13KDa protein"
                     /protein_id="AAC18571.1"
                     /db_xref="GI:3136268"
                     /translation="MSREITARANKNVPIVVGVCVVAFFVLLAFMQQKHKTHSGGDYG
                     VPTFSNGGKYRDGTRSADFNSNNHRAYGCGGSKSSVTGKVGQQLLVLALVVAVFVLFM
                     RGCWSSPEHICNGSCG"
     CDS             3307. .3513
                     /note="putative open reading frame nested within the 13
                     kDa gene"
                     /codon_start=1
                     /product="8KDa protein"
                     /protein_id="AAC18572.1"
                     /db_xref="GI:3136269"
                     /translation="MCLLLLVFVLLLSLYCWLSCNKNIRRTLVAITGFQLFRTVESIG
                     MGLGLRTLIVIIIVLMAVEVLNLV"
     CDS             3613. .4011
                     /codon_start=1
                     /product="15KDa protein"
                     /protein_id="AAC18573.1"
                     /db_xref="GI:3136270"
                     /translation="MVLVVKVDFSTIVLYIVAGVVVVSVLYSPFFSNEVKAGGYAGAI
                     FPNGGCIMDRNSFAQFGGCDIPKYVADSISRVAIKELDADIKADLNSVVAKRVVLYEG
                     LAQLCYRVFSWLVCLFMVCLMLFVWFWYHS"
     CDS             4039. .4401
                     /codon_start=1
                     /product="14KDa protein"
                     /protein_id="AAC18574.1"
                     /db_xref="GI:3136271"
                     /translation="MEKSNSIGVYVKDPITNDCRLFSVKCGNWCLFTNHVFVTYRGKN
                     DDEKVVKDTCRLHFHVKCVSCSSKVTFKANNRDHLEWLSKGFVRVNRNFSIVGACSKC
                     RGVFDSCAQQDELDNNVV"
     polyA_site      4616
BASE COUNT     1124 a    779 c   1220 g   1493 t
ORIGIN

  Query Match             61.0%;  Score 217.8;  DB 14;  Length 4616;
  Best Local Similarity   75.6%;  Pred. No. 8.2e-53;
  Matches  270;  Conservative    0;  Mismatches   87;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
           |||||||| |||||||| || ||  | ||||||||||||||||||||||||||||||||
Db    3273 ATGTCTAGAGAAATAACGGCCCGTGCTAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 3332

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120
           ||||| |||||||||||||||||||| |||||||| |||||||||||||| || |||||
Db    3333 GTTGTTGCTTTCTTTGTATTGCTGGCTTTCATGCAACAAAAACATAAGACGCACTCTGGT 3392

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
           || |||||||| || ||||| ||||| |||||||| |  ||||| || || || || ||
Db    3393 GGCGATTACGGGGTTCCAACTTTTTCGAACGGTGGAAAGTATAGGGATGGGACTAGGTCT 3452

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 240
           || || |||||||||||| |||||||||||||| || || || || |||     ||| ||
Db    3453 GCGGACTTTAATAGTAATAATCATCGTGCTTATGGCTGTGGAGGTTCTAAATCTAGTGTG 3512

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
           | | |    |||||||||||| ||  ||||| | ||  | ||    |   |  | || |
Db    3513 ACTGGCAAGGTTGGGCAGCAGCTGCTTGTGCTTGCCCTAGTAGTAGCTGTCTTTGTTTTG 3572

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
             |  | |   | |   ||||  | || || || ||||| ||||||  ||||||||||
Db    3573 TTCATGCGTGGTTGCTGGTCCTCTCCTGAGCATATTTGCAATGGTTCTTGTGGTTAA 3629


RESULT 13
LRVBABCD
LOCUS       LRVBABCD                3065 bp    RNA     linear   VRL 06-DEC-1995
DEFINITION  Lychnis ringspot virus RNA for beta-A, beta-B, beta-C, beta-D
            proteins.
ACCESSION   Z46351
VERSION     Z46351.1  GI:1107714
KEYWORDS    coat protein.
SOURCE      Lychnis ringspot virus.
  ORGANISM  Lychnis ringspot virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Hordeivirus.
REFERENCE   1  (bases 1 to 3065)
  AUTHORS   Solovyev,A.G.
  JOURNAL   Unpublished
```

```
  REMARK    (sites)
REFERENCE   2  (bases 1 to 3065)
  AUTHORS   Solovyev,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-1994) Solovyev A., Institute of Biochemistry and
            Plant Virology, Messeweg 11/12, Braunschweig, Germany
FEATURES             Location/Qualifiers
     source          1. .3065
                     /organism="Lychnis ringspot virus"
                     /db_xref="taxon:44421"
     misc_feature    1. .3065
                     /note="genomic RNA beta, partial sequence"
     CDS             157. .744
                     /function="coat protein"
                     /citation=[1]
                     /codon_start=1
                     /product="beta-A protein"
                     /protein_id="CAA86469.1"
                     /db_xref="GI:1107715"
                     /db_xref="SPTREMBL:Q83076"
                     /translation="MANLGLTVKGGGHYDEDQWNSNIIDSKVIYDWWISLDSWHSLIS
                     TLSKVQTDVVSSRRAIADLIDGVKKDLPAALDRRFAGSQGVLGAANYRAPPLFVRLDS
                     AMRANLIAIARMGDQPFLSRDQELSRNRPNTGTNQQPSNALVARDVQPLRDSALHFTY
                     DLKDLVTSDPPVFDRRSFEAKFQLNWVARAPGDNA"
     CDS             850. .2211
                     /function="virus cell-to-cell movement protein"
                     /citation=[1]
                     /codon_start=1
                     /product="beta-B protein"
                     /protein_id="CAA86470.1"
                     /db_xref="GI:1107716"
                     /db_xref="SPTREMBL:Q83077"
                     /translation="MNIFLVYNTMAIDKIVKSDNRSQKQQSSKKKKNRKHGNKVVENS
                     TNVGVPPNEGVTLRRVRSVESIDEIFTSNQRTPPAAEQLPQKQKQSKPPSVAKKHAAN
                     DKRRYEKADAFMNKIVEASRASGFQPTGQNFKRCPANLFEKCKLRALYDKHLRVLHKD
                     ACDSEKERIAAKSFLHRSLRPNVDFKVGIVSGVAGSGKSTLIRKLCSEADAMCVLANP
                     RLKETDYKGQSKTFTLQQVLLSIVPMTSDIVIVDEYTLAESAELLLLQRKLQATFLVL
                     FGDVAQGNAKTASSLEYLQFPVVFISKTSHRLGKHTAELCKKHGQAFEPGSPEEDEII
                     VADYLGAADTTEKNIAFTKETVEDLRDAGVEASLVLETQGKEYESVTLFIRESDEAAM
                     ADSHLRAVALTRHRKKLIIRAEPGVQSSFLNGELKSKTSADSHKYESSKVSYADSSSA
                     AAQ"
     CDS             2180. .2551
                     /function="virus cell-to-cell movement protein"
                     /citation=[1]
                     /codon_start=1
                     /product="beta-D protein"
                     /protein_id="CAA86471.1"
                     /db_xref="GI:1107717"
                     /db_xref="SPTREMBL:Q83078"
                     /translation="MPTQVAQRPNKYWPIVVGVGLIGLFAYLIFTNQKHATQSGDNIH
                     KFANGGSYQDGNKRINYNKNNNLAYGYKGLSNASSVDLWMLGLCIAAIAAGIYGEYLR
                     KRKHDECTTCPPDCKICGGWR"
     CDS             2406. .2879
                     /function="virus cell-to-cell movement protein"
                     /citation=[1]
                     /codon_start=1
                     /product="beta-C protein"
                     /protein_id="CAA86472.1"
                     /db_xref="GI:1107718"
                     /db_xref="SPTREMBL:Q83079"
                     /translation="MPHPLTCGCSDCVLPPLQREFTESIYANENMMSAPPAPQTAKSV
                     VGGDDLFLMMYSFLAGVLLTLLVIWSVGGTTCSNPQKASYYYQDLNKIEMEVAPGSPI
                     DPEVIKAIHHFQNFPYGRTPGLGWIDDVTTLVTVWFTRLIYLGIIFFFIWVFKNI"
BASE COUNT      889 a    597 c    681 g    898 t
ORIGIN

  Query Match             19.2%;  Score 68.6;  DB 14;  Length 3065;
  Best Local Similarity   60.8%;  Pred. No. 3.5e-09;
  Matches  132;  Conservative    0;  Mismatches   79;  Indels    6;  Gaps    1;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
           ||| | |   || || |    || |||||||||| ||  |||||| ||||||||||  ||
Db    2180 ATGCCGACTCAAGTAGCGCAGCGGCCCAATAAGTATTGGCCTATAGTTGTTGGTGTGGGT 2239

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120
             | | | | | ||||  |  |||   || |   | ||||| ||    || || || ||
Db    2240 TTGATAGGTCTTTTTGCTTATCTGATCTTTACCAATCAAAAGCACGCAACGCAATCAGGT 2299

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
           | | ||       | |  | |||| | |||||||||   |||  |||||||| | | |
Db    2300 GATAAT------ATTCACAAATTTGCCAACGGTGGTTCTTATCAAGACGGTAATAAAAGA 2353

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGT 217
             ||| ||| |  || ||| ||   |||||||| |
Db    2354 ATAAATTATAACAAAAATAATAATTTAGCTTACGGTT 2390


RESULT 14
BSBV3GNS
LOCUS       BSBV3GNS                3005 bp    RNA     linear   VRL 05-DEC-2000
DEFINITION  Beet soil-borne virus genes for 13K, 22K and 48K proteins.
ACCESSION   Z66493
VERSION     Z66493.2  GI:11595427
KEYWORDS    13K protein; 22K protein; 48K protein.
SOURCE      Beet soil-borne virus.
  ORGANISM  Beet soil-borne virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Pomovirus.
REFERENCE   1  (bases 1 to 3005)
  AUTHORS   Koenig,R., Beier,C., Commandeur,U., Luth,U., Kaufmann,A. and
            Luddecke,P.
  TITLE     Beet soil-borne virus RNA 3--a further example of the heterogeneity
            of the gene content of furovirus genomes and of triple gene
            block-carrying RNAs
  JOURNAL   Virology 216 (1), 202-207 (1996)
  MEDLINE   96187804
   PUBMED   8614988
REFERENCE   2  (bases 1 to 3005)
  AUTHORS   Koenig,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-1995) Koenig R., Biologische Bundesanstalt fuer
            Land- und Forstwirtschaft, Institut fuer
            Biochemie/Pflanzenvirologie, Messeweg 11, Braunschweig, Germany,
            D38104
  REMARK    Revised by [3]
REFERENCE   3  (bases 1 to 3005)
  AUTHORS   Koenig,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-DEC-2000) Koenig R., Biologische Bundesanstalt fuer
            Land- und Forstwirtschaft, Institut fuer
            Biochemie/Pflanzenvirologie, Messeweg 11, Braunschweig, Germany,
            D38104
COMMENT     On Dec 6, 2000 this sequence version replaced gi:1225930.
FEATURES             Location/Qualifiers
     source          1. .3005
                     /organism="Beet soil-borne virus"
                     /strain="ahlum"
                     /db_xref="taxon:46436"
     5'UTR           1. .404
     CDS             405. .1688
                     /function="Triple gene block protein, presumably cell to
                     cell movement"
                     /note="putative. NTP-binding motif present on amino acids
                     179-186"
                     /citation=[1]
                     /codon_start=1
                     /product="48K protein"
                     /protein_id="CAA91254.1"
                     /db_xref="GI:1225931"
                     /db_xref="SPTREMBL:Q65726"
                     /translation="MEKEKLQKKERNLNRRTNKKGVRRLDKKNSEEKTKEDNRKSVEE
                     DVVDGLGAHESVAKDRAATCERGKSQRQRVDQETSVKEGAGVDSKLGSDRYAGKRQLE
                     VVSRICQESGFVATGKPLKRYPEDYFLKSGLLADFDKYLSDRLDKGCNLTKSETETVL
                     KHLRQKRRPQSFLAGSVTGVPGSGKTTLLRKIQTEAGLNSVVILANERHKIRFTQLPA
                     CYTAKEILLLRTAIKYDVLLIDEYTLLQNGEILLLQRILEAKVVVLFGDRAQGNSRTA
                     DSPEWLQIPVIYSSVKSRRFGKATADFCGKQGFDFEGCDQEDEVQKLDFEGSSPETDI
                     NLALTEATIEDLKEVGIECSLVKDVQGNEYDSVSLFIREEDRAALSDPELRSVAFTRH
```

```
                RKLLIVRIPVCLMLSLFNGELNSDYRPQTNHYGKN"
     CDS        1675. .2025
                /function="Triple gene block protein, presumably cell to
                cell movement"
                /note="putative"
                /citation=[1]
                /codon_start=1
                /product="13K protein"
                /protein_id="CAA91255.1"
                /db_xref="GI:1225932"
                /db_xref="SPTREMBL:Q65727"
                /translation="MVRTNEIGARPNKYWPIVVGVFAIALFSFLTITNQKHATESGDN
                IHKFSNGGKFQDGNKRVHYNKNNPRAYNGSSSNNTFSKLLLPGLFLAAIMYAYVQYRK
                PQCTVTCRGECAHG"
     CDS        1889. .2461
                /function="triple gene block protein, presumable cell to
                cell movement"
                /note="putative"
                /citation=[1]
                /codon_start=1
                /product="22K protein"
                /protein_id="CAA91256.1"
                /db_xref="GI:1225933"
                /db_xref="SPTREMBL:Q65728"
                /translation="MDPPAIIHSQNCCCPDCSWQPSCTHTSNTESLNARSPVEENVRM
                VKGYSTDYVIMCVIVSVSLGFAIAAYFYSSGGHYDPDAAIFKQDLNEVQINFGKHPID
                PKVIEAVHHWQRAPFGRFQGESVDVSKNFFEDSVDEDDEEERAEVNKEEVTKKKSKVA
                VCNRLCSIRNSLLRLFWVITLIIIVVKCRF"
     3'UTR      2462. .3005
BASE COUNT      917 a    507 c    664 g    917 t
ORIGIN

  Query Match           17.6%;  Score 62.8;  DB 14;  Length 3005;
  Best Local Similarity  60.1%;  Pred. No. 1.7e-07;
  Matches  125;  Conservative    0;  Mismatches   77;  Indels    6;  Gaps    1;

Qy       5 CTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTG 64
           |||  || ||   |||  |||| |||||| ||  ||| || || || |||||  | ||
Db    1685 CTAATGAGATCGGCGCCAGACCTAATAAGTATTGGCCAATAGTCGTGGGTGTATTTGCAA 1744

Qy      65 TGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGGGGTG 124
           | || || ||    ||  |  |  | |   | || |||||    ||| | || || | |
Db    1745 TTGCATTGTTCAGTTTTTTAACTATAACTAATCAGAAACACGCCACAGAGTCGGGAGATA 1804

Qy     125 ATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGCTG 184
           ||       |||  |  |||||||||||||| | || |  |||||| |  |     | |
Db    1805 ATA------TCCACAAGTTTTCTAACGGTGGAAAATTTCAAGACGGCAATAAGCGTGTTC 1858

Qy     185 ATTTTAATAGTAACAATCATCGTGCTTA 212
           ||| |||||  || |||| | | |||||
Db    1859 ATTATAATAAGAATAATCCTAGAGCTTA 1886


RESULT 15
PSZBETACD
LOCUS       PSZBETACD               3591 bp ss-RNA     linear   VRL 02-AUG-1996
DEFINITION  Poa semilatent virus beta genomic segment beta-A, beta-B, beta-C
            and beta-D protein genes, complete cds.
ACCESSION   M81486
VERSION     M81486.1  GI:1478094
KEYWORDS    .
SOURCE      Poa semilatent virus cDNA to genomic RNA.
  ORGANISM  Poa semilatent virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Hordeivirus.
REFERENCE   1  (bases 1 to 3591)
  AUTHORS   Agranovsky,A.A., Karasev,A.V., Novikov,V.K., Lunina,N.A.,
            Loginov,S. and Tyulkina,L.G.
  TITLE     Poa semilatent virus, a hordeivirus having no internal polydisperse
            poly(A) in the 3' non-coding region of the RNA genome
  JOURNAL   J. Gen. Virol. 73 (Pt 8), 2085-2092 (1992)
  MEDLINE   92356087
   PUBMED   1645144
REFERENCE   2  (bases 1 to 3591)
  AUTHORS   Solovyev,A.G., Savenkov,E.I., Agranovsky,A.A. and Morozov,S.Y.
  TITLE     Comparisons of the genomic cis-elements and coding regions in RNA
            beta components of the hordeiviruses barley stripe mosaic virus,
            lychnis ringspot virus, and poa semilatent virus
  JOURNAL   Virology 219 (1), 9-18 (1996)
  MEDLINE   96204567
   PUBMED   8623558
REFERENCE   3  (bases 1 to 3591)
  AUTHORS   Solovyev,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-AUG-1993) Andrey G. Solovyev, Institute of
            Biochemistry and Plant Virology, Braunschweig, D-38104, Germany
COMMENT     On Aug 2, 1996 this sequence version replaced gi:333326.
FEATURES             Location/Qualifiers
     source          1. .3591
                     /organism="Poa semilatent virus"
                     /db_xref="taxon:12328"
                     /chromosome="RNA beta"
     CDS             97. .699
                     /function="coat protein"
                     /codon_start=1
                     /product="beta-A protein"
                     /protein_id="AAB05576.1"
                     /db_xref="GI:1478095"
                     /translation="MPNISLTTKGGGHYDEDLWFSQVVSQGLYAEWWVLSDNWEKFLD
                     NLRGVNFGVSSSRSQVSDYVAALDRDLPADVDRRFAGPRGQLGSPNYVSAPIFYRINR
                     SLVGDITRLARITDQPSNNRDVELSRVKKLSGQPTPPPGTDNQMAIRDVQSLRDGSLS
                     FSYTLINELVGAQLAAYTVRRFETEFGLQWLPVQAQGQRD"
     CDS             826. .2556
                     /function="virus cell-to-cell movement protein"
                     /codon_start=1
                     /product="beta-B-protein"
                     /protein_id="AAB05577.1"
                     /db_xref="GI:1478096"
                     /translation="MSDGLRTNHQNETPGVQSEKMELRQTTSDDTQNQADSGAKSHDT
                     NTSRVDETAKTTDVKESPSGESLSGVSAVSGGTDKGEKSAVRKEAGESKESPKSVPSD
                     GGEVKHVDTETKAKSKRKKKNKKTPKEGTSKTTSESSSAKNVESKESKKQTKPKAVSP
                     SSDTSKIASEVGSAKKATKKESKKQTKDKGSAEDLNANTKLKAKASEQKGPTIPGTSA
                     EASRIDLLQSSTAEKFTKNDVRRTALVNEFVAQIIHKFCIEQGFEPTGRQYMRARANLF
                     ELVGLRNLYMEHLKKTAAKACNFTKDRIRRKLFLTSNHKPSVDFLVGIVSGVPGCGKS
                     TLVRKLLDSPISCYVALANPATERDYRGTSNVMTLDDLLLAKVPMSSDLLIVDEYTLA
                     ESAEILLLQRRVRSSLVLLVGDVAQGRSNNASNIEYLTLPVIYRKITSHRIGEETAKA
                     CSKQGNRIRSAGRKDKLILVDYEGETEETEKNLAFTEETRDDVKDCGYDCSLVSEVQG
                     LEYESVTLFLRNTDRAAASDHHKRTVSMTRHKSLLIIRAEQEIGQPFLMGELSVSSKR
                     PSNAHVYSSE"
     CDS             2559. .2969
                     /function="virus cell-to-cell movement protein"
                     /codon_start=1
                     /product="beta-D protein"
                     /protein_id="AAB05578.1"
                     /db_xref="GI:1478097"
                     /translation="MPAKTTSVGSRPNKYWPIVAGIGVVGLFAYLIFANQKHSTESGD
                     NIHKFANGGSYRDGSKCISYHRNNPFAYGNASSPGMLLPTLFTIIGIVSYLWATRGEI
                     MGGNSPLLGHNCGEECAGECVRGLDRQELPPNSV"
     CDS             2773. .3261
                     /function="virus cell-to-cell movement protein"
                     /codon_start=1
                     /product="beta-C protein"
                     /protein_id="AAB05579.1"
                     /db_xref="GI:1478098"
                     /translation="MAMPHLLGCSCPLCSQSSESFPIYGPPEERLWEETPPSSATIVE
                     RNAPENVFVAWIDRNYLLILCSLALLSVSLAYIYFTSGNNNPYVKGGYFYQDLNSVEV
                     RFGEHPVDPKVIASIHHWQKNPFGVSPLWEGLNNLVSSLVSALKFYVGLLFLLLLIII
                     FK"
BASE COUNT      1049 a    718 c    842 g    982 t
ORIGIN

  Query Match           16.2%;  Score 57.8;  DB 14;  Length 3591;
  Best Local Similarity  60.0%;  Pred. No. 5.1e-06;
  Matches  117;  Conservative    0;  Mismatches   72;  Indels    6;  Gaps    1;

Qy      21 TCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATT 80
           | |||| |||||| ||  ||| ||||| | |||| |  |||| || | ||| ||||  |
```

```
Db   2588 TAGACCTAATAAGTATTGGCCAATTGTAGCTGGTATAGGTGTAGTAGGTTTATTTGCTTA 2647

Qy     81 GCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGGGGTGATTACGGAGTCCCAAC 140

Db   2648 CTTAATTTTTGCTAATCAGAAACATTCTACTGAGTCTGGAGATAATA------TCCACAA 2701

Qy    141 ATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAA 200

Db   2702 ATTTGCCAACGGTGGGAGTTATCGAGACGGTTCTAAGTGTATTTCTTACCATCGAAACAA 2761

Qy    201 TCATCGTGCTTACGG 215

Db   2762 CCCTTTCGCCTATGG 2776
```

Search completed: March 12, 2003, 09:31:30
Job time : 1179 secs

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: March 11, 2003, 14:24:19 ; Search time 175 Seconds
(without alignments)
4594.077 Million cell updates/sec

Title: US-09-936-011-1
Perfect score: 357
Sequence: 1 atgtctagggaaataaccgc.....gtaatggtgcttgtggttaa 357

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	357	100.0	357	21	AAA75877	DNA encoding a P13
	2	347.4	97.3	357	21	AAA75878	DNA encoding a P13
	3	337.8	94.6	739	22	AAF26890	Beet necrotic yell
	4	337.8	94.6	4609	22	AAF26889	Beet necrotic yell
	5	325	91.0	739	22	AAF26891	Beet necrotic yell
	6	53.8	15.1	507	24	ABK86954	Potato mop-top vir
	7	45.8	12.8	507	24	ABK86955	Mutant potato mop-
	8	36.6	10.3	4039	22	AAH34822	Human colon cancer
	9	36.6	10.3	4039	24	ABL90346	Human polynucleoti
	10	36.6	10.3	4040	21	AAC77977	Human cancer assoc
	11	36.6	10.3	5544	24	ABL61761	Colon adenocarcino
	12	36.6	10.3	5549	23	AAS87055	DNA encoding novel
	13	35.6	10.0	1668	24	ABK40083	Human chemically p
	14	35.6	10.0	6725	24	ABL33209	Human immune syste
	15	35.6	10.0	6725	24	ABL34555	Human metastasis a
	16	35.2	9.9	3001	21	AAH51765	Chromosome 13q31-q
c	17	35	9.8	759	19	AAV28675	Ripening banana pu
	18	34.6	9.7	2772	23	ABL25868	Drosophila melanog
	19	34.4	9.6	1143	23	AAS77369	DNA encoding novel
	20	34.4	9.6	2263	23	AAS88931	DNA encoding novel
c	21	34.2	9.6	2328	23	AAS84696	DNA encoding novel
	22	34.2	9.6	2494	24	ABK31539	Signal transductio
c	23	34.2	9.6	5296	22	AAK77661	Human immune/haema
	24	34	9.5	7434	24	AAD28386	Human chemically t
c	25	33.8	9.5	406	24	ABN75893	Human ORF840 cDNA,
	26	33.8	9.5	6531	24	ABL32640	Human immune syste
	27	33.6	9.4	1110	21	AAC81556	Human growth facto
	28	33.6	9.4	4590	22	AAH24065	Yeast AOD9604-asso
c	29	33.2	9.3	1683	21	AAC37459	Arabidopsis thalia
	30	33.2	9.3	5971	24	ABL32383	Human immune syste
	31	33	9.2	14147	22	AAS46743	Tumour suppressor
	32	33	9.2	14147	24	ABK33955	Human DNA for stag
	33	32.8	9.2	6535	24	ABL32937	Human immune syste
	34	32.6	9.1	5986	24	AAS61433	Human gene regulat
	35	32.6	9.1	5986	24	ABK31499	Signal transductio
	36	32.4	9.1	487	23	ABV51629	Human prostate exp
c	37	32.4	9.1	2527	15	AAQ56756	Chitinase 1 gene.
c	38	32.4	9.1	2585	24	ABL87903	Human ovarian canc
c	39	32.4	9.1	3297	23	AAS91437	DNA encoding novel
c	40	32.4	9.1	29376	23	ABL08834	Drosophila melanog
c	41	32.2	9.0	234	22	ABA71836	Human foetal liver
c	42	32.2	9.0	234	22	AAK20212	Human brain expres
c	43	32.2	9.0	234	22	AAK46281	Human bone marrow
c	44	32.2	9.0	234	22	AAI52180	Probe #20866 used
c	45	32.2	9.0	234	24	ABS20595	Human genome-deriv

ALIGNMENTS

RESULT 1
AAA75877
ID AAA75877 standard; DNA; 357 BP.
XX
AC AAA75877;
XX
DT 22-JAN-2001 (first entry)
XX
DE DNA encoding a P13 protein of Beet necrotic yellow vein mosaic virus.
XX
KW P13 protein; triple gene block 2; TGB2; resistance; group I virus; ss.
XX
OS Beet necrotic yellow vein mosaic virus.
XX
FH Key Location/Qualifiers
FT CDS 1..357
FT /*tag= a
FT /product= "P13 protein"
XX
PN WO200055301-A2.
XX
PD 21-SEP-2000.
XX
PF 07-MAR-2000; 2000WO-EP02176.
XX
PR 12-MAR-1999; 99EP-0200773.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Jonard G, Lauber E, Guilley H, Richards K;
XX
DR WPI; 2000-602114/57.

```
DR   P-PSDB; AAB18801.
XX
PT   Inducing viral resistance into a plant or a plant cell, especially beet
PT   necrotic yellow vein virus-resistance into a sugar beet plant or cell
PT   by transforming the plant cell with triple gene block of the virus -
XX
PS   Disclosure; Fig 1; 27pp; English.
XX
CC   The present sequence encodes a P13 protein of Beet necrotic yellow
CC   vein mosaic virus. The sequence is a triple gene block 2 (TGB2)
CC   sequence. It is used for inducing resistance to a group I virus. The
CC   method is useful for inducing resistance to a plant cell or plants
CC   such as sugar beet, potato, barley or peanut against group I virus
CC   such as hordeiviruses, benyviruses, pecluviruses and pomoviruses,
CC   preferably barley stripe mosaic virus, potato mop top virus, peanut
CC   clump virus and the beet soil-borne virus, more preferably beet
CC   necrotic yellow vein virus (BNYVV).
XX
SQ   Sequence 357 BP; 89 A; 58 C; 93 G; 117 T; 0 other;

  Query Match             100.0%;  Score 357;  DB 21;  Length 357;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-98;
  Matches  357;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
Db       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120
Db      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
Db     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240
Db     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
Db     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
Db     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357


RESULT 2
AAA75878
ID   AAA75878 standard; DNA; 357 BP.
XX
AC   AAA75878;
XX
DT   22-JAN-2001  (first entry)
XX
DE   DNA encoding a P13 variant of Beet necrotic yellow vein mosaic virus.
XX
KW   P13 protein; triple gene block 2; TGB2; resistance; group I virus; ss.
XX
OS   Beet necrotic yellow vein mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..357
FT                   /*tag=  a
FT                   /product= "P13 protein"
XX
PN   WO200055301-A2.
XX
PD   21-SEP-2000.
XX
PF   07-MAR-2000; 2000WO-EP02176.
XX
PR   12-MAR-1999;   99EP-0200773.
XX
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Jonard G,  Lauber E,  Guilley H,  Richards K;
XX
DR   WPI; 2000-602114/57.
DR   P-PSDB; AAB18802.
XX
PT   Inducing viral resistance into a plant or a plant cell, especially beet
PT   necrotic yellow vein virus-resistance into a sugar beet plant or cell
PT   by transforming the plant cell with triple gene block of the virus -
XX
PS   Disclosure; Fig 2; 27pp; English.
XX
CC   The present sequence encodes a P13 protein variant of Beet necrotic
CC   yellow vein mosaic virus. The P13 sequence is a triple gene block 2
CC   (TGB2) sequence. It is used for inducing resistance to a group I virus.
CC   The method is useful for inducing resistance to a plant cell or plants
CC   such as sugar beet, potato, barley or peanut against group I virus
CC   such as hordeiviruses, benyviruses, pecluviruses and pomoviruses,
CC   preferably barley stripe mosaic virus, potato mop top virus, peanut
CC   clump virus and the beet soil-borne virus, more preferably beet
CC   necrotic yellow vein virus (BNYVV).
XX
SQ   Sequence 357 BP; 84 A; 60 C; 96 G; 117 T; 0 other;

  Query Match             97.3%;  Score 347.4;  DB 21;  Length 357;
  Best Local Similarity   98.3%;  Pred. No. 4.9e-95;
  Matches  351;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
Db       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120
Db      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGG 120

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
Db     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240
Db     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
Db     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
Db     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357


RESULT 3
AAF26890
ID   AAF26890 standard; DNA; 739 BP.
XX
AC   AAF26890;
XX
DT   09-APR-2001  (first entry)
XX
DE   Beet necrotic yellow vein virus promoter sequence SEQ ID NO:2.
XX
KW   Beet necrotic yellow vein virus; BNYVV; transformed plant;
KW   Rhizomania disease-resistant plant; promoter; ds.
XX
OS   Beet necrotic yellow vein mosaic virus.
XX
PN   JP2000312540-A.
```

```
XX
PD   14-NOV-2000.
XX
PF   28-APR-1999;   99JP-0122628.
XX
PR   28-APR-1999;   99JP-0122628.
XX
PA   (HOKK-) HOKKAIDO PREFECTURE.
PA   (HOKK-) HOKKAIDO TENSAI KYOKAI SH.
XX
DR   WPI; 2001-054202/07.
XX
PT   A Rhizomania disease-resistant plant  -
XX
PS   Claim 14; Page 9; 11pp; Japanese.
XX
CC   The present invention describes a method for producing a transformed
CC   plant in which resistance against beet necrotic yellow vein virus
CC   (BNYVV) is given by transforming expressably a gene derived from BNYVV
CC   genome or a DNA corresponding to its part or a DNA substantially same
CC   as it in a plant genome. The vector structure can be used for
CC   transforming a plant or a plant cell having BNYVV resistance. The
CC   present sequence represents a wild type BNYVV promoter nucleotide
CC   sequence for use in the method of the invention.
XX
SQ   Sequence 739 BP; 184 A; 114 C; 189 G; 252 T; 0 other;

  Query Match             94.6%;  Score 337.8;  DB 22;  Length 739;
  Best Local Similarity   96.6%;  Pred. No. 5e-92;
  Matches  345;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
Db       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120
Db      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
Db     121 GGCGATTACGGAGTCCCAACATTTTCTAACGGTGGTAAATACAGAGACGGTACAAGGTCA 180

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 240
Db     181 GCTGATTTTAATAGTAATAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 240

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
Db     241 AGTAGTCGAGTCGGGCAGCAACTTGTTGTGTTAGCTATTGTGTCTGTGTTAATAGTATCA 300

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
Db     301 CTGTTACAACGATTAAGATCTCCACCAGAACATATTTGTAATGGTGCTTGTGGTTAA 357


RESULT 4
AAF26889
ID   AAF26889 standard; DNA; 4609 BP.
XX
AC   AAF26889;
XX
DT   09-APR-2001  (first entry)
XX
DE   Beet necrotic yellow vein virus RNA-2 nucleotide sequence SEQ ID NO:1.
XX
KW   Beet necrotic yellow vein virus; BNYVV; transformed plant;
KW   Rhizomania disease-resistant plant; ds.
XX
OS   Beet necrotic yellow vein mosaic virus.
XX
PN   JP2000312540-A.
XX
PD   14-NOV-2000.
XX
PF   28-APR-1999;   99JP-0122628.
XX
PR   28-APR-1999;   99JP-0122628.
XX
PA   (HOKK-) HOKKAIDO PREFECTURE.
PA   (HOKK-) HOKKAIDO TENSAI KYOKAI SH.
XX
DR   WPI; 2001-054202/07.
XX
PT   A Rhizomania disease-resistant plant  -
XX
PS   Claim 13; Page 7-9; 11pp; Japanese.
XX
CC   The present invention describes a method for producing a transformed
CC   plant in which resistance against beet necrotic yellow vein virus
CC   (BNYVV) is given by transforming expressably a gene derived from BNYVV
CC   genome or a DNA corresponding to its part or a DNA substantially same
CC   as it in a plant genome. The vector structure can be used for
CC   transforming a plant or a plant cell having BNYVV resistance. The
CC   present sequence represents a specifically claimed BNYVV nucleotide
CC   sequence for use in the method of the invention.
XX
SQ   Sequence 4609 BP; 1187 A; 765 C; 1221 G; 1436 T; 0 other;

  Query Match             94.6%;  Score 337.8;  DB 22;  Length 4609;
  Best Local Similarity   96.6%;  Pred. No. 9.7e-92;
  Matches  345;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
Db    3284 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 3343

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120
Db    3344 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 3403

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
Db    3404 GGCGATTACGGAGTCCCAACATTTTCTAACGGTGGTAAATACAGAGACGGTACAAGGTCA 3463

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 240
Db    3464 GCTGATTTTAATAGTAATAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 3523

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
Db    3524 AGTAGTCGAGTCGGGCAGCAACTTGTTGTGTTAGCTATTGTGTCTGTGTTAATAGTATCA 3583

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
Db    3584 CTGTTACAACGATTAAGATCTCCACCAGAACATATTTGTAATGGTGCTTGTGGTTAA 3640


RESULT 5
AAF26891
ID   AAF26891 standard; DNA; 739 BP.
XX
AC   AAF26891;
XX
DT   09-APR-2001  (first entry)
XX
DE   Beet necrotic yellow vein virus mutant promoter sequence SEQ ID NO:3.
XX
KW   Beet necrotic yellow vein virus; BNYVV; transformed plant;
KW   Rhizomania disease-resistant plant; promoter; ds.
XX
OS   Beet necrotic yellow vein mosaic virus.
OS   Synthetic.
XX
PN   JP2000312540-A.
XX
```

```
PD   14-NOV-2000.
XX
PF   28-APR-1999;   99JP-0122628.
XX
PR   28-APR-1999;   99JP-0122628.
XX
PA   (HOKK-) HOKKAIDO PREFECTURE.
PA   (HOKK-) HOKKAIDO TENSAI KYOKAI SH.
XX
DR   WPI; 2001-054202/07.
XX
PT   A Rhizomania disease-resistant plant  -
XX
PS   Claim 15; Page 9; 11pp; Japanese.
XX
CC   The present invention describes a method for producing a transformed
CC   plant in which resistance against beet necrotic yellow vein virus
CC   (BNYVV) is given by transforming expressably a gene derived from BNYVV
CC   genome or a DNA corresponding to its part or a DNA substantially same
CC   as it in a plant genome. The vector structure can be used for
CC   transforming a plant or a plant cell having BNYVV resistance. The
CC   present sequence represents a mutant BNYVV promoter nucleotide
CC   sequence for use in the method of the invention.
XX
SQ   Sequence 739 BP; 182 A; 117 C; 187 G; 253 T; 0 other;

  Query Match             91.0%;  Score 325;  DB 22;  Length 739;
  Best Local Similarity   94.4%;  Pred. No. 3.7e-88;
  Matches  337;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     .  1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
           || |||||||||||||||||||||||||| || || ||  ||||  |   |||| |||
Db     121 GGCGATTACGGAGTCCCAACATTTTCTAACGCTGCTAAATTCAGAGCTGCAGCAAGGTCA 180

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 240
           ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db     181 GCTGATTTTAATAGTAATAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 240

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
           |||||||||| |||||||||||| |||||||||||||||| |||||||||||||| |||
Db     241 AGTAGTCGAGTCGGGCAGCAACTTGTTGTGTTAGCTATTGTGTCTGTGTTAATAGTATCA 300

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
           || |||||||||||||| |||||||||||||| ||||||||||||||||||||||||
Db     301 CTGTTACAACGATTAAGATCTCCACCAGAACATATTTGTAATGGTGCTTGTGGTTAA 357


RESULT 6
ABK86954
ID   ABK86954 standard; cDNA; 507 BP.
XX
AC   ABK86954;
XX
DT   24-SEP-2002  (first entry)
XX
DE   Potato mop-top virus (PMTV) triple-gene-block (TGB)-2 cDNA.
XX
KW   Potato; antiviral; triple-gene-block 2; TGB-2; PMTV;
KW   potato mop-top virus; Pomovirus; spraing; dwarfing; mopping;
KW   chlorotic chevron; necrotic chevron; blotching; tuber cracking;
KW   necrotic conchoidal layer; viral movement; transgenic; molecular marker;
KW   resistance; agriculture; gene; ss.
XX
OS   Potato mop-top virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             37..396
FT                   /*tag=  a
FT                   /product= "TGB-2"
XX
PN   WO200250281-A1.
XX
PD   27-JUN-2002.
XX
PF   18-DEC-2001; 2001WO-SE02805.
XX
PR   21-DEC-2000; 2000SE-0004755.
XX
PA   (PLAN-) PLANT SCI SWEDEN AB.
XX
PI   Melander M,  Lee M;
XX
DR   WPI; 2002-508804/54.
DR   P-PSDB; AAU79576.
XX
PT   Nucleic acid molecule for producing plants with increased resistance
PT   against infection by potato mop-top virus comprises a structural gene
PT   encoding a triple-gene-block 2 polypeptide from the virus -
XX
PS   Disclosure; Page 31; 36pp; English.
XX
CC   The invention discloses a nucleic acid molecule comprising a promoter
CC   operably linked to a structural gene encoding a triple-gene-block 2
CC   (TGB-2) polypeptide from potato mop-top virus (PMTV) which differs from
CC   the wild-type polypeptide at a number of residues. The potato crop
CC   (Solanum tuberosum) is susceptible to PMTV which belongs to the genus
CC   Pomovirus. The main symptoms of the disease are spraing, dwarfing
CC   (mopping), chlorotic and necrotic chevrons, blotching and tuber cracking
CC   and necrotic conchoidal layers. PMTV has a tripartite RNA genome, of
CC   which RNA 2 encodes for four proteins in a triple-gene-block and all
CC   three seem to be responsible for long-distance viral movement within the
CC   plant. The mutant TGB-2 is useful for producing transgenic plants, in
CC   particular, the potato plant, with an increased resistance against
CC   infection by PMTV. Portions of the mutant TGB-2 are useful as molecular
CC   markers for detecting a transgenic plant and for detecting a food product
CC   from the transgenic plant. The advantage the mutant TGB-2 brings is that
CC   it is able to render potato plants resistant by using solely a single
CC   gene which is easy to follow and maintain during a breeding program. The
CC   transgenic plants can be grown even in fields infected by PMTV, which
CC   increases the economic value of the yield obtained from such an infected
CC   field. The sequence presented is the wild-type TGB-2 cDNA from PMTV.
XX
SQ   Sequence 507 BP; 143 A; 94 C; 106 G; 164 T; 0 other;

  Query Match             15.1%;  Score 53.8;  DB 24;  Length 507;
  Best Local Similarity   57.5%;  Pred. No. 2.6e-06;
  Matches  119;  Conservative    0;  Mismatches   82;  Indels    6;  Gaps    1;

Qy       6 TAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGT 65
           ||  |||||    || ||||| |||||  ||  |||  | |||||  |  ||   ||   |
Db      48 TAACGAAATTGGAGCGCGACCAAATAAATATTGGCCGGTAGTTGTGGCAGTAGTTGCAAT 107

Qy      66 GGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGGGGTGA 125
             | | || |  ||  |  |  | |   | ||||||||    || || || || | | |
Db     108 TTGTCTTTTCGGTTTTTTAACAGTTACCAATCAAAAACACGCTACTCAATCAGGTGATAA 167

Qy     126 TTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGA 185
           |       | |  | |||| ||||||||||  | || || ||||||  | |       | |
Db     168 TA------TACATAAATTTGCTAACGGTGGCCTGTACAGGGACGGTTCTAAGAGTATTAA 221

Qy     186 TTTTAATAGTAACAATCATCGTGCTTA 212
             | |||| |||| ||||   | |||||
Db     222 GTATAATTGTAATAATCCCAGAGCTTA 248


RESULT 7
ABK86955
ID   ABK86955 standard; cDNA; 507 BP.
```

```
XX
AC   ABK86955;
XX
DT   24-SEP-2002  (first entry)
XX
DE   Mutant potato mop-top virus (PMTV) triple-gene-block (TGB)-2 cDNA.
XX
KW   Potato; antiviral; triple-gene-block 2; TGB-2; PMTV; 13kDmut;
KW   potato mop-top virus; Pomovirus; spraing; dwarfing; mopping;
KW   chlorotic chevron; necrotic chevron; blotching; tuber cracking;
KW   necrotic conchoidal layer; viral movement; transgenic; molecular marker;
KW   resistance; agriculture; gene; ss; mutant.
XX
OS   Potato mop-top virus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             37..396
FT                   /*tag= a
FT                   /product= "TGB-2 mutant (13kDmut)"
FT   mutation        replace(188,G)
FT                   /*tag= b
FT   mutation        replace(191,G)
FT                   /*tag= c
FT   mutation        replace(193,C)
FT                   /*tag= d
FT   mutation        replace(194,T)
FT                   /*tag= e
FT   mutation        replace(195,G)
FT                   /*tag= f
FT   mutation        replace(197,A)
FT                   /*tag= g
FT   mutation        replace(204,C)
FT                   /*tag= h
FT   mutation        replace(206,G)
FT                   /*tag= i
FT   mutation        replace(208,T)
FT                   /*tag= j
XX
PN   WO200250281-A1.
XX
PD   27-JUN-2002.
XX
PF   18-DEC-2001; 2001WO-SE02805.
XX
PR   21-DEC-2000; 2000SE-0004755.
XX
PA   (PLAN-) PLANT SCI SWEDEN AB.
XX
PI   Melander M,  Lee M;
XX
DR   WPI; 2002-508804/54.
DR   P-PSDB; AAU79577.
XX
PT   Nucleic acid molecule for producing plants with increased resistance
PT   against infection by potato mop-top virus comprises a structural gene
PT   encoding a triple-gene-block 2 polypeptide from the virus -
XX
PS   Example 3; Page 32; 36pp; English.
XX
CC   The invention discloses a nucleic acid molecule comprising a promoter
CC   operably linked to a structural gene encoding a triple-gene-block 2
CC   (TGB-2) polypeptide from potato mop-top virus (PMTV) which differs from
CC   the wild-type polypeptide at a number of residues. The potato crop
CC   (Solanum tuberosum) is susceptible to PMTV which belongs to the genus
CC   Pomovirus. The main symptoms of the disease are spraing, dwarfing
CC   (mopping), chlorotic and necrotic chevrons, blotching and tuber cracking
CC   and necrotic conchoidal layers. PMTV has a tripartite RNA genome, of
CC   which RNA 2 encodes for four proteins in a triple-gene-block and all
CC   three seem to be responsible for long-distance viral movement within the
CC   plant. The mutant TGB-2 is useful for producing transgenic plants, in
CC   particular, the potato plant, with an increased resistance against
CC   infection by PMTV. Portions of the mutant TGB-2 are useful as molecular
CC   markers for detecting a transgenic plant and for detecting a food product
CC   from the transgenic plant. The advantage the mutant TGB-2 brings is that
CC   it is able to render potato plants resistant by using solely a single
CC   gene which is easy to follow and maintain during a breeding program. The
CC   transgenic plants can be grown even in fields infected by PMTV, which
CC   increases the economic value of the yield obtained from such an infected
CC   field. The sequence presented is a mutant TGB-2 (13kDmut) cDNA from PMTV.
XX
SQ   Sequence 507 BP; 145 A; 95 C; 104 G; 163 T; 0 other;

  Query Match             12.8%;  Score 45.8;  DB 24;  Length 507;
  Best Local Similarity   55.1%;  Pred. No. 0.00067;
  Matches  114;  Conservative    0;  Mismatches   87;  Indels    6;  Gaps    1;

Qy       6 TAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGT 65
           ||  |||||    || ||||| |||||  ||  |||  | ||||| |  ||   ||  |
Db      48 TAACGAAATTGGAGCGCGACCAAATAAATATTGGCCGGTAGTTGTGGCAGTAGTTGCAAT 107

Qy      66 GGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGGGGTGA 125
             | | || |  ||  |  |  | |   | ||||||||    || || || || | | |
Db     108 TTGTCTTTTCGGTTTTTTAACAGTTACCAATCAAAAACACGCTACTCAATCAGGTGATAA 167

Qy     126 TTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGA 185
           |       | |  | |||| |||||| ||    ||  || || | | | |       | |
Db     168 TA------TACATAAATTTGCTAACGCTGCCGAATTCAGGGAAGCTGCTAAGAGTATTAA 221

Qy     186 TTTTAATAGTAACAATCATCGTGCTTA 212
             | |||| |||| ||||   | |||||
Db     222 GTATAATTGTAATAATCCCAGAGCTTA 248


RESULT 8
AAH34822
ID   AAH34822 standard; cDNA; 4039 BP.
XX
AC   AAH34822;
XX
DT   03-SEP-2001  (first entry)
XX
DE   Human colon cancer antigen encoding cDNA SEQ ID NO:1904.
XX
KW   Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW   colorectal carcinoma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200122920-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26524.
XX
PR   29-SEP-1999;   99US-0157137.
PR   03-NOV-1999;   99US-0163280.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Barash SC,  Birse CE,  Rosen CA;
XX
DR   WPI; 2001-235357/24.
DR   P-PSDB; AAG75417.
XX
PT   Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT   useful for preventing, diagnosing and/or treating colorectal cancers -
XX
PS   Claim 1; Page 3415-3417; 9803pp; English.
XX
CC   AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC   cancer-associated nucleic acid molecules (N) and proteins (P), where
CC   the proteins are collectively known as colon cancer antigens. The colon
CC   cancer antigens have cytostatic activity and can be used in gene
CC   therapy and vaccine production. N and P may be used in the prevention,
CC   diagnosis and treatment of diseases associated with inappropriate P
```

CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ Sequence 4039 BP; 1174 A; 790 C; 789 G; 1283 T; 3 other;

Query Match 10.3%; Score 36.6; DB 22; Length 4039;
Best Local Similarity 49.7%; Pred. No. 0.85;
Matches 93; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

Qy 123 TGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGC 182
Db 585 TGACCCCCTGGCACCAACACCCGCCAGTACTGTGACTTCCAAAGCCAGAGCCACATGTGC 644

Qy 183 TGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAG 242
Db 645 TCATCAAACTTGCATTAAGCAGTTGGCGGGAGATGGCTGTGGAGCTGGGGGTTTAAGTGA 704

Qy 243 TAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCACT 302
Db 705 TGGTTCTCTTTTGCTCCCTCTTYTGAGGGTAAAGCTACTGTCTTTCTTAAGAGTGTATTT 764

Qy 303 ATTACAA 309
Db 765 ATGCCAA 771

RESULT 9
ABL90346
ID ABL90346 standard; cDNA; 4039 BP.
XX
AC ABL90346;
XX
DT 24-MAY-2002 (first entry)
XX
DE Human polynucleotide SEQ ID NO 908.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200190304-A2.
XX
PD 29-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-US16450.
XX
PR 19-MAY-2000; 2000US-205515P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
DR WPI; 2002-122018/16.
DR P-PSDB; ABB89937.
XX
PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders -
XX
PS Claim 4; SEQ ID NO 908; 2081pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 4039 BP; 1174 A; 790 C; 789 G; 1283 T; 3 other;

Query Match 10.3%; Score 36.6; DB 24; Length 4039;
Best Local Similarity 49.7%; Pred. No. 0.85;
Matches 93; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

Qy 123 TGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGC 182
Db 585 TGACCCCCTGGCACCAACACCCGCCAGTACTGTGACTTCCAAAGCCAGAGCCACATGTGC 644

Qy 183 TGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAG 242
Db 645 TCATCAAACTTGCATTAAGCAGTTGGCGGGAGATGGCTGTGGAGCTGGGGGTTTAAGTGA 704

Qy 243 TAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCACT 302
Db 705 TGGTTCTCTTTTGCTCCCTCTTYTGAGGGTAAAGCTACTGTCTTTCTTAAGAGTGTATTT 764

Qy 303 ATTACAA 309
Db 765 ATGCCAA 771

RESULT 10
AAC77977
ID AAC77977 standard; cDNA; 4040 BP.
XX
AC AAC77977;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human cancer associated gene sequence SEQ ID NO:371.
XX
KW Human; cancer associated gene; cancer antigen; detection; cancer;
KW diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW allergic reaction; graft versus host disease; organ rejection;
KW haemostatic; thrombolytic; cardiovascular disorder; infection;
KW neurological disease; drug screening; ss.
XX
OS Homo sapiens.
XX
PN WO200055350-A1.
XX
PD 21-SEP-2000.
XX

```
PF   08-MAR-2000; 2000WO-US05882.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-587533/55.
DR   P-PSDB; AAB43768.
XX
PT   Novel isolated nucleic acids comprising sequences encoding peptides
PT   useful for treating or diagnosing e.g. cancer -
XX
PS   Claim 1; Page 918-920; 2352pp; English.
XX
CC   AAC77607 to AAC78448 encode the human cancer associated proteins given
CC   in AAB43398 to AAB44239. The proteins can have activities based on the
CC   tissues and cells the genes are expressed in. Example of activities
CC   include: cytostatic; proliferative; vulnerary; immunomodulator;
CC   antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC   antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC   dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC   nootropic; vasotropic; antipsoriatic and antiangiogenic. The
CC   polynucleotides and polypeptides can be used for preventing, treating or
CC   ameliorating medical conditions and diagnosing pathological conditions.
CC   Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC   the present invention may be used to treat immune disorders by activating
CC   or inhibiting the proliferation, differentiation or mobilisation of
CC   immune cells, to treat disorders of haematopoietic cells, autoimmune
CC   disorders, allergic reactions, graft versus host disease and organ
CC   rejection, modulate haemostatic or thrombolytic activity, modulate
CC   inflammation, cancers, cardiovascular disorders, neurological disease and
CC   bacterial or viral infections. The peptides, nucleotides, antibodies,
CC   agonists and antagonists may be also be used in drug screens. AAC78449 to
CC   AAC78457 and AAB44240 represent sequences used in the exemplification of
CC   the present invention.
XX
SQ   Sequence 4040 BP; 1173 A; 789 C; 790 G; 1281 T; 7 other;

  Query Match             10.3%;  Score 36.6;  DB 21;  Length 4040;
  Best Local Similarity   49.7%;  Pred. No. 0.85;
  Matches   93;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy      123 TGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGC 182
            |||   |   |  ||||||    | |    ||  |  |  | || | |  | | ||  ||
Db      586 TGACCCCCTGGCACCAACACCCGCCAGTACTGTGACTTCCAAAGCCAGAGCCACATGTGC 645

Qy      183 TGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAG 242
            | ||   | | | |  || ||    ||    |   || |||| ||||||||      |
Db      646 TCATCAAACTTGCATTAAGCAGTTGGCGGGAGATGGCTGTGGAGCTGGGGGTTTAAGTGA 705

Qy      243 TAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCACT 302
            | ||    ||  ||  |  ||| |   | ||    |  | ||| | |||| |||||   |
Db      706 TGGTTCTCTTTTGCTCCCTCTTYTGAGGGTAAAGCTACTGTCTTTCTTAAGAGTGTATTT 765

Qy      303 ATTACAA 309
            ||  |||
Db      766 ATGCCAA 772


RESULT 11
ABL61761
ID   ABL61761 standard; DNA; 5544 BP.
XX
AC   ABL61761;
XX
DT   15-MAY-2002  (first entry)
XX
DE   Colon adenocarcinoma related gene sequence SEQ ID NO:98.
XX
KW   Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW   stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW   cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW   gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200194629-A2.
XX
PD   13-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US10838.
XX
PR   05-JUN-2000; 2000US-209473P.
PR   05-JUN-2000; 2000US-209531P.
PR   18-SEP-2000; 2000US-233133P.
PR   18-SEP-2000; 2000US-233617P.
PR   20-SEP-2000; 2000US-234009P.
PR   20-SEP-2000; 2000US-234034P.
PR   20-SEP-2000; 2000US-234052P.
PR   22-SEP-2000; 2000US-234509P.
PR   22-SEP-2000; 2000US-234567P.
PR   25-SEP-2000; 2000US-234923P.
PR   25-SEP-2000; 2000US-234924P.
PR   25-SEP-2000; 2000US-235077P.
PR   25-SEP-2000; 2000US-235082P.
PR   25-SEP-2000; 2000US-235134P.
PR   25-SEP-2000; 2000US-235280P.
PR   26-SEP-2000; 2000US-235637P.
PR   26-SEP-2000; 2000US-235638P.
PR   27-SEP-2000; 2000US-235711P.
PR   27-SEP-2000; 2000US-235720P.
PR   27-SEP-2000; 2000US-235840P.
PR   27-SEP-2000; 2000US-235863P.
PR   28-SEP-2000; 2000US-236028P.
PR   28-SEP-2000; 2000US-236032P.
PR   28-SEP-2000; 2000US-236033P.
PR   28-SEP-2000; 2000US-236034P.
PR   28-SEP-2000; 2000US-236109P.
PR   28-SEP-2000; 2000US-236111P.
PR   29-SEP-2000; 2000US-236842P.
PR   29-SEP-2000; 2000US-236891P.
PR   02-OCT-2000; 2000US-237172P.
PR   02-OCT-2000; 2000US-237173P.
PR   02-OCT-2000; 2000US-237278P.
PR   02-OCT-2000; 2000US-237294P.
PR   02-OCT-2000; 2000US-237295P.
PR   02-OCT-2000; 2000US-237316P.
PR   03-OCT-2000; 2000US-237425P.
PR   03-OCT-2000; 2000US-237598P.
PR   03-OCT-2000; 2000US-237604P.
PR   03-OCT-2000; 2000US-237606P.
PR   03-OCT-2000; 2000US-237608P.
PR   01-NOV-2000; 2000US-244867P.
PR   01-NOV-2000; 2000US-245084P.
XX
PA   (AVAL-) AVALON PHARM.
XX
PI   Young PE,  Augustus M,  Carter KC,  Ebner R,  Endress G,  Horrigan S;
PI   Soppet DR,  Weaver Z;
XX
DR   WPI; 2002-188264/24.
XX
PT   Screening for anti-neoplastic agent involves exposing cells to a
PT   chemical agent to be tested for anti-neoplastic activity, and
PT   determining a change in expression of a gene of a signature gene set -
XX
PS   Claim 1; SEQ ID 98; 44pp; English.
XX
CC   The present invention describes a method (M1) for screening for an
CC   anti-neoplastic agent. The method involves exposing cells to a chemical
CC   agent to be tested for anti-neoplastic activity, determining a change in
CC   expression of at least one gene (I) of a signature gene set, where (I)
CC   comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC   to ABL70110), or is at least 95% identical to (S), where a change in
```

```
CC   expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC   activity and can be used in gene therapy. M1 can be used for screening
CC   an anti-neoplastic agent, and can be used for producing a product which
CC   is the data collected with respect to the anti-neoplastic agent as a
CC   result of M1, and the data is sufficient to convey the chemical
CC   structure and/or properties of the agent. M1 can be used in the
CC   treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC   oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC   adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC   infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC   carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ   Sequence 5544 BP; 1504 A; 1281 C; 1157 G; 1602 T; 0 other;

  Query Match             10.3%;  Score 36.6;  DB 24;  Length 5544;
  Best Local Similarity   49.7%;  Pred. No. 0.95;
  Matches   93;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy     123 TGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGC 182
           |||   |   |  ||||||    | |    ||  |  |  | || | |  | | ||  ||
Db    2106 TGACCCCCTGGCACCAACACCCGCCAGTACTGTGACTTCCAAAGCCAGAGCCACATGTGC 2165

Qy     183 TGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAG 242
           | ||   | | | |  || ||    ||     |   || ||||  ||||||||     |
Db    2166 TCATCAAACTTGCATTAAGCAGTTGGCGGGAGATGGCTGTGGAGCTGGGGGTTTAAGTGA 2225

Qy     243 TAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCACT 302
           | ||    ||  ||  |  ||| |   | ||    |  | ||| | |||| |||||   |
Db    2226 TGGTTCTCTTTTGCTCCCTCTTTTGAGGGTAAAGCTACTGTCTTTCTTAAGAGTGTATTT 2285

Qy     303 ATTACAA 309
           ||  |||
Db    2286 ATGCCAA 2292


RESULT 12
AAS87055
ID   AAS87055 standard; cDNA; 5549 BP.
XX
AC   AAS87055;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #22859.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG22868.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity
XX
PS   Claim 1; SEQ ID No 22859; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 5549 BP; 1510 A; 1285 C; 1154 G; 1600 T; 0 other;

  Query Match             10.3%;  Score 36.6;  DB 23;  Length 5549;
  Best Local Similarity   49.7%;  Pred. No. 0.95;
  Matches   93;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy     123 TGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGC 182
           |||   |   |  ||||||    | |    ||  |  |  | || | |  | | ||  ||
Db    2106 TGACCCCCTGGCACCAACACCCGCCAGTACTGTGACTTCCAAAGCCAGAGCCACATGTGC 2165

Qy     183 TGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAG 242
           | ||   | | | |  || ||    ||     |   || ||||  ||||||||     |
Db    2166 TCATCAAACTTGCATTAAGCAGTTGGCGGGAGATGGCTGTGGAGCTGGGGGTTTAAGTGA 2225

Qy     243 TAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCACT 302
           | ||    ||  ||  |  ||| |   | ||    |  | ||| | |||| |||||   |
Db    2226 TGGTTCTCTTTTGCTCCCTCTTTTGAGGGTAAAGCTACTGTCTTTCTTAAGAGTGTATTT 2285

Qy     303 ATTACAA 309
           ||  |||
Db    2286 ATGCCAA 2292


RESULT 13
ABK40083
ID   ABK40083 standard; DNA; 1668 BP.
XX
AC   ABK40083;
XX
DT   21-MAY-2002  (first entry)
XX
DE   Human chemically pretreated gene sequence #83 strand 1.
XX
KW   Human; ds; bisulphite treatment; CpG; DNA methylation; cancer; tumour;
KW   cytostatic; ALDH6; CYP11A; CYP11B1; CYP3A3; DPYD; EPHX2; OCLN; TXNRD1;
KW   UGT8; MRP; pharmacogenomics; SNP; single nucleotide polymorphism.
XX
OS   Homo sapiens.
XX
PN   WO200202806-A2.
XX
PD   10-JAN-2002.
XX
PF   29-JUN-2001; 2001WO-EP07470.
XX
PR   30-JUN-2000; 2000DE-1032529.
PR   01-SEP-2000; 2000DE-1043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
```

```
XX
DR   WPI; 2002-154757/20.
XX
PT   New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,
PT   useful for detecting cytosine methylation state of genes associated
PT   with pharmacogenomics and for therapy of diseases e.g. cancer -
XX
PS   Claim 1; SEQ ID No 165; 24pp; English.
XX
CC   The invention relates to a nucleic acid comprising a sequence at
CC   least 18 bases in length of a segment of the chemically pretreated DNA
CC   of genes associated with pharmacogenomics according to one of the
CC   sequences of the genes ALDH6 (NM_000693), CYP11A (NM_000781), CYP11B1
CC   (NM_000497), CYP3A3 (NM_000776 and NM_017460), DPYD (NM_000110), EPHX2
CC   (NM_001979), OCLN (NM_002538), TXNRD1 (NM_003330), UGT8 (NM_003360),
CC   MRP (NM_004996, NM_019900, NM_019901, NM_019902, NM_019862, NM_019898,
CC   NM_019899) and their complementary sequences, or a sequence (S1) chosen
CC   from 87 sequences and their complements. The chemical pretreatment
CC   is bisulphite treatment to convert cytosines (but not methyl-cytosines)
CC   into uracils. Also included are an oligomer (II) in particular an
CC   oligonucleotide or a peptide nucleic acid (PNA)-oligomer, comprising in
CC   each case at least one base sequence having a length of 9 nucleotides
CC   which hybridises to or is identical to a chemically pretreated DNA of
CC   genes associated with pharmacogenomics and their complements, arranged in
CC   an array for analysing diseases associated with the methylation state
CC   (CpG) and/or detecting SNPs (single nucleotide polymorphisms)
CC   of the 87 sequences. The oligomers may also be used as PCR primers.
CC   The set of 87 nucleic acids and their complements is useful for diagnosis
CC   and therapy of solid tumours and cancer. The present sequence
CC   represents one the 87 DNA sequences or its complement.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1668 BP; 480 A; 15 C; 359 G; 814 T; 0 other;

  Query Match             10.0%;  Score 35.6;  DB 24;  Length 1668;
  Best Local Similarity   51.2%;  Pred. No. 1.2;
  Matches   83;  Conservative    0;  Mismatches   79;  Indels    0;  Gaps    0;

Qy     141 ATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAA 200
           ||||| ||  |  ||||||| |||  | | |     |   |  |  ||||| | |||  ||
Db     387 ATTTTTTATTGTGGGTATATTTAGGTAAGTTTTTTGTGTAAGTTTTTTTATTTGTATAAA 446

Qy     201 TCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAGTAGTCGAGTTGGGCAGCA 260
             ||   |  || |  |  || |  | || || | | || |  | | |    ||   |
Db     447 ATATTTAGTGTAAGTATTTGGGGTTTTTGTGGATTGGGTAATGATTTGGAAGGGTTGGAG 506

Qy     261 ACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCACT 302
             || ||| |     | || ||  ||| |   || ||| | |
Db     507 GTTTTTTGGGGATTTTTTTTTTTATTGTTTGTTTAATGTAAGT 548


RESULT 14
ABL33209
ID   ABL33209 standard; DNA; 6725 BP.
XX
AC   ABL33209;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human immune system associated gene SEQ ID NO: 1182.
XX
KW   Human; immune system disease; cytosine methylation; antiasthmatic;
KW   antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW   neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW   antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW   antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW   acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW   neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW   gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200200928-A2.
XX
PD   03-JAN-2002.
XX
PF   02-JUL-2001; 2001WO-EP07537.
XX
PR   30-JUN-2000; 2000DE-1032529.
PR   01-SEP-2000; 2000DE-1043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-130909/17.
XX
PT   Nucleic acid comprising fragment of chemically modified gene, useful
PT   for diagnosis and treatment of diseases associated with abnormal
PT   cytosine methylation -
XX
PS   Claim 1; SEQ ID NO 1182; 32pp + Sequence Listing; German.
XX
CC   The present invention provides a number of human immune system associated
CC   genes which are modified by the methylation of cytosines. The sequences
CC   can be used in the diagnosis and treatment of immune system disorders,
CC   including eye diseases such as retinopathy, neovascular glaucoma and
CC   macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC   leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC   rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC   diseases. The present sequence is a gene of the invention.
XX
SQ   Sequence 6725 BP; 2074 A; 53 C; 1429 G; 3169 T; 0 other;

  Query Match             10.0%;  Score 35.6;  DB 24;  Length 6725;
  Best Local Similarity   48.1%;  Pred. No. 2;
  Matches  101;  Conservative    0;  Mismatches  109;  Indels    0;  Gaps    0;

Qy       3 GTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGT 62
           || | | |||||  |   |  ||  || |   |        |||    ||||| | |
Db    3443 GTATGGTGAAATTTCGTTTTTACGAAAAATAGAAAAATTAGTTGGGTGTGGTGGTGGGTG 3502

Qy      63 TGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGGGG 122
           | |    || |   ||||    |  ||| | | || | || | |   ||  |     || ||
Db    3503 TTTATAATTTTAGTTATTTGGGATGTTGAGGTAGGAGAATCGTTTGAATCGAGGAGGCGG 3562

Qy     123 TGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGC 182
           | ||   | |   | | ||| | | |  ||  | ||  | | ||  |     ||
Db    3563 AGGTTGTAGTGAGTCGAGATTGTGTTATTGTATTTTAGTTTGGGAAAGAGAGTGAGATTT 3622

Qy     183 TGATTTTAATAGTAACAATCATCGTGCTTA 212
           || ||| |||| ||| ||| ||  |  | |
Db    3623 TGTTTTAAATAATAATAATAATAATAATAA 3652


RESULT 15
ABL34555
ID   ABL34555 standard; DNA; 6725 BP.
XX
AC   ABL34555;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human metastasis associated gene SEQ ID NO: 108.
XX
KW   Metastasis associated gene; cytostatic; gene therapy; cancer;
KW   cytosine methylation; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200177376-A2.
XX
PD   18-OCT-2001.
```

XX
PF 06-APR-2001; 2001WO-EP03970.
XX
PR 06-APR-2000; 2000DE-1019058.
PR 07-APR-2000; 2000DE-1019173.
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-010922/01.
XX
PT New nucleic acid derived from chemically treated metastasis genes,
PT useful for diagnosis of cancers by analysis of cytosine methylation,
PT also for treatment -
XX
PS Claim 1; SEQ ID NO 108; 23pp + Sequence Listing; English.
XX
CC The present invention provides a number of human metastasis associated
CC genes which are modified by cytosine methylation. The sequences can be
CC used in the diagnosis and treatment of cancer. The present sequence is
CC one of the genes of the invention.
XX
SQ Sequence 6725 BP; 2074 A; 53 C; 1429 G; 3169 T; 0 other;

Query Match 10.0%; Score 35.6; DB 24; Length 6725;
Best Local Similarity 48.1%; Pred. No. 2;
Matches 101; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

```
Qy    3 GTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGT 62
        || | | |||||  |   |  ||  || |   |       |||    ||||| | |
Db 3443 GTATGGTGAAATTTCGTTTTTACGAAAAATAGAAAAATTAGTTGGGTGTGGTGGTGGGTG 3502

Qy   63 TGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGGGG 122
        | |    || |   ||||   |  ||| | | || | || | |  ||  |     || ||
Db 3503 TTTATAATTTTAGTTATTTGGGATGTTGAGGTAGGAGAATCGTTTGAATCGAGGAGGCGG 3562

Qy  123 TGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGC 182
         | ||   | |   | | ||| | | |  ||  | ||  | | ||  |     ||
Db 3563 AGGTTGTAGTGAGTCGAGATTGTGTTATTGTATTTTAGTTTGGGAAAGAGAGTGAGATTT 3622

Qy  183 TGATTTTAATAGTAACAATCATCGTGCTTA 212
        || ||| |||| ||| ||| ||  |  | |
Db 3623 TGTTTTAAATAATAATAATAATAATAATAA 3652
```

Search completed: March 12, 2003, 08:17:17
Job time : 189 secs

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: March 11, 2003, 10:26:12 ; Search time 53.5 Seconds
(without alignments)
2046.423 Million cell updates/sec

Title: US-09-936-011-1
Perfect score: 357
Sequence: 1 atgtctagggaaataaccgc.....gtaatggtgcttgtggttaa 357

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	30.8	8.6	2848	3	US-08-539-205A-3	Sequence 3, Appli
	2	30.4	8.5	799	4	US-09-149-476-288	Sequence 288, App
	3	30.4	8.5	1847	4	US-09-149-476-149	Sequence 149, App
	4	30.4	8.5	1847	4	US-09-149-476-287	Sequence 287, App
	5	30.4	8.5	9595	3	US-09-014-416-4	Sequence 4, Appli
	6	29.6	8.3	4285	4	US-09-410-464-1	Sequence 1, Appli
	7	29.2	8.2	11309	4	US-08-961-527-108	Sequence 108, App
	8	29	8.1	1370	4	US-09-026-408-12	Sequence 12, Appl
	9	29	8.1	1371	4	US-09-026-408-1	Sequence 1, Appli
c	10	29	8.1	1857	4	US-09-134-001C-332	Sequence 332, App
	11	28.8	8.1	2362	1	US-08-920-812-5	Sequence 5, Appli
	12	28.8	8.1	2362	1	US-08-920-827-5	Sequence 5, Appli
	13	28.8	8.1	2362	1	US-08-921-177-5	Sequence 5, Appli
	14	28.8	8.1	2362	1	US-08-362-577C-5	Sequence 5, Appli
	15	28.8	8.1	2362	2	US-08-920-828-5	Sequence 5, Appli
c	16	28.8	8.1	5883	1	US-08-404-445-1	Sequence 1, Appli
	17	28.8	8.1	6060	5	PCT-US96-09430-7	Sequence 7, Appli
c	18	28.8	8.1	8442	4	US-09-272-032-6	Sequence 6, Appli
	19	28.6	8.0	875	4	US-08-936-165A-125	Sequence 125, App
	20	28.6	8.0	1506	2	US-08-663-566A-8	Sequence 8, Appli
	21	28.6	8.0	1506	2	US-08-023-610-8	Sequence 8, Appli
	22	28.6	8.0	1506	2	US-08-288-065A-8	Sequence 8, Appli
	23	28.6	8.0	1506	2	US-08-362-240A-8	Sequence 8, Appli
	24	28.6	8.0	1506	4	US-08-804-372A-6	Sequence 6, Appli
	25	28.6	8.0	1506	5	PCT-US95-10245-8	Sequence 8, Appli
	26	28.6	8.0	1825	4	US-09-071-035-463	Sequence 463, App
	27	28.6	8.0	2601	3	US-09-039-773A-2	Sequence 2, Appli
	28	28.6	8.0	6168	4	US-09-071-035-457	Sequence 457, App
	29	28.6	8.0	6168	4	US-09-071-035-461	Sequence 461, App
	30	28.6	8.0	6168	4	US-09-071-035-465	Sequence 465, App
c	31	28.4	8.0	660	4	US-09-549-804C-1	Sequence 1, Appli
	32	28.4	8.0	2910	4	US-09-206-942-31	Sequence 31, Appl
	33	28.4	8.0	2928	4	US-09-206-942-29	Sequence 29, Appl
	34	28.4	8.0	3454	4	US-08-961-527-207	Sequence 207, App
c	35	28.4	8.0	16569	4	US-09-097-889-2	Sequence 2, Appli
c	36	28.4	8.0	16569	4	US-09-377-856-1	Sequence 1, Appli
c	37	28.4	8.0	16569	4	US-09-302-681-2	Sequence 2, Appli
	38	28.4	8.0	18596	4	US-09-318-448-11	Sequence 11, Appl
	39	28.2	7.9	1170	5	PCT-US96-05320A-900	Sequence 900, App
c	40	28.2	7.9	5511	3	US-08-928-361B-2	Sequence 2, Appli
c	41	28.2	7.9	7334	3	US-08-928-361B-1	Sequence 1, Appli
c	42	28	7.8	549	1	US-08-196-945-1	Sequence 1, Appli
	43	28	7.8	846	2	US-08-743-637B-189	Sequence 189, App
	44	28	7.8	7766	4	US-09-125-619-3	Sequence 3, Appli
c	45	27.8	7.8	2087	4	US-09-097-199-83	Sequence 83, Appl

ALIGNMENTS

RESULT 1
US-08-539-205A-3
; Sequence 3, Application US/08539205A
; Patent No. 6001619
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/539,205A
; FILING DATE: 04-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2848 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 238..2535
US-08-539-205A-3

Query Match 8.6%; Score 30.8; DB 3; Length 2848;
Best Local Similarity 50.0%; Pred. No. 3.1;
Matches 77; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

Qy 5 CTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTG 64

```
Db     377 CTATCAAAAAGACGTTAAACCCTTATTGGAATGAGACTTTTGAAGTTAATGTTACAGATA 436

Qy      65 TGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGGGGTG 124
           | |||   | | || ||||   ||| |||||  ||||| | |       ||
Db     437 ATAGCACCATTGCGATTCAGGTGTTCGATCAGAAAAAATTTAAGAAAAAAGGCCAAGGCT 496

Qy     125 ATTACGGAGTCCCAACATTTTCTAACGGTGGTAT 158
             |   |||||   ||   ||  |   || | | |
Db     497 TTCTAGGAGTGATAAATCTTCGTGTTGGAGATGT 530


RESULT 2
US-09-149-476-288
; Sequence 288, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 186 Human Secreted proteins
;  FILE REFERENCE: PZ002P1
;  CURRENT APPLICATION NUMBER: US/09/149,476
;  CURRENT FILING DATE: 1998-09-08
;  EARLIER APPLICATION NUMBER: PCT/US98/04493
;  EARLIER FILING DATE: 1998-03-06
;  EARLIER APPLICATION NUMBER: 60/040,162
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,333
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/038,621
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,626
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,334
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,336
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,163
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/047,600
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,615
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,597
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,502
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,633
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,583
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,617
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,618
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,503
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,592
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,581
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,584
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,500
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,587
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,492
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,598
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,613
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,582
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,596
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,612
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,632
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,601
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/043,580
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,568
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,314
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,569
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,311
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,671
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,674
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,669
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,312
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,313
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,672
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,315
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/048,974
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/056,886
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,877
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,889
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,893
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,630
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,878
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,662
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,872
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,882
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,637
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,903
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,888
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,879
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,880
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,894
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,911
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,636
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,874
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,910
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,864
```

; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match 8.5%; Score 30.4; DB 4; Length 799;
Best Local Similarity 61.2%; Pred. No. 2.3;
Matches 49; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

```
Qy      34 AATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATG 93
           || | |||   ||| ||  |||||   || ||  ||| | | | ||| |  | || |||
Db     678 AACATACCTTAAGTTTTTTTTGTTTTGTTTTTGTTTTTTTGTTTTTTGTTTTCCTTTATG 737

Qy      94 CAGCAAAAACATAAGACACA 113
              || ||||| | || |  ||
Db     738 AAGAAAAAATAAAATAGTCA 757
```

RESULT 3
US-09-149-476-149
; Sequence 149, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: PZ002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612

; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match 8.5%; Score 30.4; DB 4; Length 1847;
Best Local Similarity 61.2%; Pred. No. 3.4;
Matches 49; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

```
Qy      34 AATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATG 93
           ||  | |||   ||| ||  |||||   || ||  ||| | | | ||| |  | || |||
Db    1724 AACATACCTTAAGTTTTTTTGTTTTGTTTTTGTTTTTTTGTTTTTTGTTTTCCTTTATG 1783
```

```
Qy      94 CAGCAAAAACATAAGACACA 113
           || ||||| | || |  ||
Db    1784 AAGAAAAAATAAAATAGTCA 1803


RESULT 4
US-09-149-476-287
; Sequence 287, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 186 Human Secreted proteins
;  FILE REFERENCE: PZ002P1
;  CURRENT APPLICATION NUMBER: US/09/149,476
;  CURRENT FILING DATE: 1998-09-08
;  EARLIER APPLICATION NUMBER: PCT/US98/04493
;  EARLIER FILING DATE: 1998-03-06
;  EARLIER APPLICATION NUMBER: 60/040,162
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,333
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/038,621
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,626
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,334
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,336
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,163
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/047,600
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,615
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,597
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,502
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,633
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,583
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,617
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,618
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,503
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,592
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,581
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,584
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,500
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,587
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,492
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,598
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,613
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,582
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,596
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,612
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,632
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,601
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/043,580
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,568
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,314
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,569
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,311
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,671
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,674
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,669
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,312
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,313
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,672
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,315
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/048,974
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/056,886
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,877
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,889
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,893
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,630
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,878
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,662
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,872
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,882
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,637
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,903
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,888
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,879
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,880
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,894
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,911
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,636
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,874
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,910
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,864
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,631
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,845
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,892
```

```
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/057,761
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/047,595
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,599
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,588
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,585
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,586
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,590
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,594
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,589
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,593
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,614
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/043,578
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,576
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/047,501
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/043,670
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/056,632
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,664
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,876
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,881
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,909
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,875
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,862
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,887
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,908
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/048,964
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/057,650
;  EARLIER FILING DATE: 1997-09-05
;  EARLIER APPLICATION NUMBER: 60/056,884
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/057,669
;  EARLIER FILING DATE: 1997-09-05
;  EARLIER APPLICATION NUMBER: 60/049,610
;  EARLIER FILING DATE: 1997-06-13
;  EARLIER APPLICATION NUMBER: 60/061,060
;  EARLIER FILING DATE: 1997-10-02

  Query Match             8.5%;  Score 30.4;  DB 4;  Length 1847;
  Best Local Similarity   61.2%;  Pred. No. 3.4;
  Matches   49;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy      34 AATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATG 93
           ||  | |||   ||| ||  |||||   || ||  ||| | | | ||| |  | || |||
Db    1734 AACATACCTTAAGTTTTTTTGTTTTGTTTTTGTTTTTTGTTTTTGTTTCCTTTATG 1793

Qy      94 CAGCAAAAACATAAGACACA 113
             || ||||| | || |  ||
Db    1794 AAGAAAAAATAAAATAGTCA 1813
```

```
RESULT 5
US-09-014-416-4
; Sequence 4, Application US/09014416
; Patent No. 6153421
; GENERAL INFORMATION:
;  APPLICANT: Yanagi, Masayuki
;  APPLICANT: Bukh, Jens
;  APPLICANT: Emerson, Susanne U.
;  APPLICANT: Purcell, Robert H.
;  TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
;  TITLE OF INVENTION: USES THEREOF
;  FILE REFERENCE: 20264276
;  CURRENT APPLICATION NUMBER: US/09/014,416
;  CURRENT FILING DATE: 1998-01-27
;  EARLIER APPLICATION NUMBER: US 60/053,062
;  EARLIER FILING DATE: 1997-07-18
;  NUMBER OF SEQ ID NOS: 65
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 9595
;   TYPE: DNA
;   ORGANISM: Hepatitis C virus
US-09-014-416-4

  Query Match             8.5%;  Score 30.4;  DB 3;  Length 9595;
  Best Local Similarity   61.2%;  Pred. No. 7.2;
  Matches   49;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
           |||    ||||  ||||| ||| |  |  ||||    | ||| || || ||  | ||| |
Db    9371 ATGAACGGGGAGCTAACCACTCCAGGCCTTAAGCCATTTCCTGTTTTTTTTTTTTTTTT 9430

Qy      61 GTTGTGGCTTTCTTTGTATT 80
             || |   |||||||  | ||
Db    9431 TTTTTTTTTTCTTTTTTT 9450
```

```
RESULT 6
US-09-410-464-1
; Sequence 1, Application US/09410464
; Patent No. 6395892
; GENERAL INFORMATION:
;  APPLICANT: Strauss et al.
;  TITLE OF INVENTION: Floral homeotic genes for manipulation of flowering in
;  TITLE OF INVENTION: poplar and other plant species.
;  FILE REFERENCE: 53375
;  CURRENT APPLICATION NUMBER: US/09/410,464
;  CURRENT FILING DATE: 1999-10-01
;  EARLIER APPLICATION NUMBER: 09/287,700
;  EARLIER FILING DATE: 1999-04-06
;  EARLIER APPLICATION NUMBER: 60/080,851
;  EARLIER FILING DATE: 1998-04-06
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 4285
;   TYPE: DNA
;   ORGANISM: Populus balsamifera subsp. trichocarpa
US-09-410-464-1

  Query Match             8.3%;  Score 29.6;  DB 4;  Length 4285;
  Best Local Similarity   54.6%;  Pred. No. 9;
  Matches   59;  Conservative    0;  Mismatches   49;  Indels    0;  Gaps    0;

Qy     175 AGATCAGCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGT 234
           | ||||| | | || || || || || || | |||    || |  || |||     |     |
Db    2540 AAATCAGGTAACTTCAAAAGAAATAACCTTCGCATATATGCATGTGGTTATGGTTTTTAT 2599

Qy     235 AGCGTTAGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTT 282
             |   ||  ||  | ||| | ||| ||| |||  | ||  | || ||
Db    2600 GGGAATATCTGTAAATTTGTGGAGCTACTAATTAAGGTATTTGTTTTT 2647
```

```
RESULT 7
US-08-961-527-108
; Sequence 108, Application US/08961527
; Patent No. 6420135
;  GENERAL INFORMATION:
;    APPLICANT:  Charles Kunsch
;    TITLE OF INVENTION:  Streptococcus pneumoniae Polynucleotides and Sequences
;    NUMBER OF SEQUENCES:  391
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/961,527
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brookes, A. Anders
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB340P1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  108:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11309 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-961-527-108

  Query Match             8.2%;  Score 29.2;  DB 4;  Length 11309;
  Best Local Similarity   51.5%;  Pred. No. 19;
  Matches   67;  Conservative    0;  Mismatches   63;  Indels    0;  Gaps    0;

Qy      25 CCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTG 84
           | |||  | ||| | |||||| ||     ||     ||||   |||| | | |  ||
Db    7011 CGCAAGGAAAATATTCCTATTATTCCCCATGAAACGGTTGCTTATTTCCGTTTTCTTATG 7070

Qy      85 GCGTTCATGCAGCAAAAACATAAGACACATTCTGGGGGTGATTACGGAGTCCCAACATTT 144
           |    |||||||| ||| |||      |   ||||  | |  |||  |  || |  |
Db    7071 GAAACCATGCAGCCTAAAAATATTCTGGAAATTGGGACGGCTATCGGTTTTTCAGCTCTC 7130

Qy     145 TCTAACGGTG 154
           |  |  | ||
Db    7131 TTGATGGCTG 7140


RESULT 8
US-09-026-408-12
; Sequence 12, Application US/09026408
; Patent No. 6303338
;  GENERAL INFORMATION:
;    APPLICANT:  Ni et al.
;    TITLE OF INVENTION:  PANCREAS-DERIVED PLASMINOGEN ACTIVATOR
;    TITLE OF INVENTION:  INHIBITOR
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;      STREET:  1100 NEW YORK AVENUE, SUITE 600
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/026,408
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/934,011
;      FILING DATE:  15-AUG-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/024,056
;      FILING DATE:  16-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  STEFFE, ERIC K.
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1488.0300002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-2600
;      TELEFAX:  202-371-2540
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1370 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  67..1281
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  67..120
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  121..1281
US-09-026-408-12

  Query Match             8.1%;  Score 29;  DB 4;  Length 1370;
  Best Local Similarity   54.1%;  Pred. No. 8.4;
  Matches   59;  Conservative    0;  Mismatches   50;  Indels    0;  Gaps    0;

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
           || |||| |  ||| ||  | ||| |||||    ||  | || |||| ||       ||||
Db      57 AGAAGTCAAAATGGACACAATCTTCTTGTGGAGTCTTCTATTGCTGTTTTTTGGAAGTCA 116

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTT 349
           |  |    | || ||   |  | | |  ||||| | |||  |||
Db     117 AGCCTCAAGATGCTCAGCTCAAAAAAATACCGAATTTGCAGTGGATCTT 165


RESULT 9
US-09-026-408-1
; Sequence 1, Application US/09026408
; Patent No. 6303338
;  GENERAL INFORMATION:
;    APPLICANT:  Ni et al.
;    TITLE OF INVENTION:  PANCREAS-DERIVED PLASMINOGEN ACTIVATOR
;    TITLE OF INVENTION:  INHIBITOR
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;      STREET:  1100 NEW YORK AVENUE, SUITE 600
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3934
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/026,408
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/934,011
;      FILING DATE:  15-AUG-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/024,056
;      FILING DATE:  16-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  STEFFE, ERIC K.
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1488.0300002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-2600
;      TELEFAX:  202-371-2540
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1371 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  both
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  67..1242
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  109..1242
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  67..108
US-09-026-408-1

  Query Match              8.1%;  Score 29;  DB 4;  Length 1371;
  Best Local Similarity   54.1%;  Pred. No. 8.4;
  Matches   59;  Conservative    0;  Mismatches   50;  Indels    0;  Gaps    0;

Qy    241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
          || |||| |  ||| ||  | ||| |||||    ||  | || |||| ||      ||||
Db     57 AGAAGTCAAAATGGACACAATCTTCTTGTGGAGTCTTCTATTGCTGTTTTTTGGAAGTCA 116

Qy    301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTT 349
             |  |     | || ||   |  | | |  ||||| | |||  |||
Db    117 AGCCTCAAGATGCTCAGCTCAAAAAAATACCGAATTTGCAGTGGATCTT 165


RESULT 10
US-09-134-001C-332/c
; Sequence 332, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 332
;  LENGTH: 1857
;  TYPE: DNA
;  ORGANISM: Staphylococcus epidermidis
US-09-134-001C-332

  Query Match              8.1%;  Score 29;  DB 4;  Length 1857;
  Best Local Similarity   51.1%;  Pred. No. 9.6;
  Matches   68;  Conservative    0;  Mismatches   65;  Indels    0;  Gaps    0;

Qy     71 TCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGGGGTGATTACG 130
          | |  |||||| |    | ||  |  ||   ||||| ||   |  | | | |   | |
Db    260 TGTCCGTATTGACGAATATTATCGATTAAGCGCATAACACGTTTAATTGTGCTATCTCCC 201

Qy    131 GAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTA 190
          |  |   |||     | | | ||||||  | |||   || ||  || |   |   |||||
Db    200 GCTTGTGAACCACCCCCATCAGTGGTAGGTTTAGCTTCGATATGAGCTTCACCTGTTTTA 141

Qy    191 ATAGTAACAATCA 203
          || |||   | ||
Db    140 ATTGTACTGAACA 128


RESULT 11
US-08-920-812-5
; Sequence 5, Application US/08920812
; Patent No. 5763188
;  GENERAL INFORMATION:
;    APPLICANT:  Ohno, Tsuneya
;    APPLICANT:  Matsuhisa, Akio
;    APPLICANT:  Uehara, Hirotsugu
;    APPLICANT:  Eda, Soji
;    TITLE OF INVENTION:  Probe for Diagnosing Infectious Disease
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/920,812
;      FILING DATE:  29-AUG-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/362,577
;      FILING DATE:  27-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  19036/32420
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2362 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Staphylococcus epidermidis
;      STRAIN:  Clinical Isolate SE-3
US-08-920-812-5

  Query Match              8.1%;  Score 28.8;  DB 1;  Length 2362;
  Best Local Similarity   51.1%;  Pred. No. 12;
  Matches   93;  Conservative    0;  Mismatches   87;  Indels    2;  Gaps    1;
```

```
Qy      21 TCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATT 80
           ||   | ||  |  ||||    |   | || | |    || |||  | |||  |||||
Db    1074 TCATACAAAAGAAGATGTTGAGAAAATAGTAGATTCAAGTATTGACGGTTTACTTGTAGG 1133

Qy      81 GCTGGCGTTCATGCAGCAAAA--ACATAAGACACATTCTGGGGGTGATTACGGAGTCCCA 138
           |  ||| || ||| |   |||  || |||| ||  || | |    | |||  |       |
Db    1134 GGAGGCATTAATGAAAACAAATGACTTAAGTCAGTTTTTTGCCTAGTTTAAAGTTAAAGA 1193

Qy     139 ACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAAC 198
           | | | ||||     | || || | | |    || ||    |  |||| |||| |  |
Db    1194 AGAATCTCTATGATAGTTAAATTTTGTGGTTTTAAAACCGAAAGTGATATTAAGAAAATT 1253

Qy     199 AA 200
           ||
Db    1254 AA 1255
```

RESULT 12
US-08-920-827-5
; Sequence 5, Application US/08920827
; Patent No. 5770375
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,827
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2362 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus epidermidis
; STRAIN: Clinical Isolate SE-3
US-08-920-827-5

Query Match 8.1%; Score 28.8; DB 1; Length 2362;
Best Local Similarity 51.1%; Pred. No. 12;
Matches 93; Conservative 0; Mismatches 87; Indels 2; Gaps 1;

```
Qy      21 TCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATT 80
           ||   | ||  |  ||||    |   | || | |    || |||  | |||  |||||
Db    1074 TCATACAAAAGAAGATGTTGAGAAAATAGTAGATTCAAGTATTGACGGTTTACTTGTAGG 1133

Qy      81 GCTGGCGTTCATGCAGCAAAA--ACATAAGACACATTCTGGGGGTGATTACGGAGTCCCA 138
           |  ||| || ||| |   |||  || |||| ||  || | |    | |||  |       |
Db    1134 GGAGGCATTAATGAAAACAAATGACTTAAGTCAGTTTTTTGCCTAGTTTAAAGTTAAAGA 1193

Qy     139 ACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAAC 198
           | | | ||||     | || || | | |    || ||    |  |||| |||| |  |
Db    1194 AGAATCTCTATGATAGTTAAATTTTGTGGTTTTAAAACCGAAAGTGATATTAAGAAAATT 1253

Qy     199 AA 200
           ||
Db    1254 AA 1255
```

RESULT 13
US-08-921-177-5
; Sequence 5, Application US/08921177
; Patent No. 5798211
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/921,177
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2362 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus epidermidis
; STRAIN: Clinical Isolate SE-3
US-08-921-177-5

Query Match 8.1%; Score 28.8; DB 1; Length 2362;
Best Local Similarity 51.1%; Pred. No. 12;
Matches 93; Conservative 0; Mismatches 87; Indels 2; Gaps 1;

```
Qy      21 TCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATT 80
           ||   | ||  |  ||||    |   | || | |    || |||  | |||  |||||
Db    1074 TCATACAAAAGAAGATGTTGAGAAAATAGTAGATTCAAGTATTGACGGTTTACTTGTAGG 1133
```

```
Qy      81 GCTGGCGTTCATGCAGCAAAA--ACATAAGACACATTCTGGGGGTGATTACGGAGTCCCA 138
           |  ||| || ||| |   |||  || |||| ||  || | |    | |||  |      |
Db    1134 GGAGGCATTAATGAAAACAAATGACTTAAGTCAGTTTTTTGCCTAGTTTAAAGTTAAAGA 1193

Qy     139 ACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAAC 198
           | | | ||||     | || || | | |    || ||    |  |||| |||| |  |
Db    1194 AGAATCTCTATGATAGTTAAATTTTGTGGTTTTAAAACCGAAAGTGATATTAAGAAAATT 1253

Qy     199 AA 200
           ||
Db    1254 AA 1255


RESULT 14
US-08-362-577C-5
; Sequence 5, Application US/08362577C
; Patent No. 5807673
;  GENERAL INFORMATION:
;    APPLICANT:  Ohno, Tsuneya
;    APPLICANT:  Matsuhisa, Akio
;    APPLICANT:  Uehara, Hirotsugu
;    APPLICANT:  Eda, Soji
;    TITLE OF INVENTION:  Probe for Diagnosing Infectious Disease
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/362,577C
;      FILING DATE:  27-MAR-1995
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  19036/32420
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2362 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Staphylococcus epidermidis
;      STRAIN:  Clinical Isolate SE-3
US-08-362-577C-5

  Query Match             8.1%;  Score 28.8;  DB 1;  Length 2362;
  Best Local Similarity   51.1%;  Pred. No. 12;
  Matches   93;  Conservative    0;  Mismatches   87;  Indels    2;  Gaps    1;

Qy      21 TCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATT 80
           ||   | ||  |  ||||    |   | || | |    || |||  | |||  |||||
Db    1074 TCATACAAAAGAAGATGTTGAGAAAATAGTAGATTCAAGTATTGACGGTTTACTTGTAGG 1133

Qy      81 GCTGGCGTTCATGCAGCAAAA--ACATAAGACACATTCTGGGGGTGATTACGGAGTCCCA 138
           |  ||| || ||| |   |||  || |||| ||  || | |    | |||  |      |
Db    1134 GGAGGCATTAATGAAAACAAATGACTTAAGTCAGTTTTTTGCCTAGTTTAAAGTTAAAGA 1193

Qy     139 ACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAAC 198
           | | | ||||     | || || | | |    || ||    |  |||| |||| |  |
Db    1194 AGAATCTCTATGATAGTTAAATTTTGTGGTTTTAAAACCGAAAGTGATATTAAGAAAATT 1253

Qy     199 AA 200
           ||
Db    1254 AA 1255


RESULT 15
US-08-920-828-5
; Sequence 5, Application US/08920828
; Patent No. 5853998
;  GENERAL INFORMATION:
;    APPLICANT:  Ohno, Tsuneya
;    APPLICANT:  Matsuhisa, Akio
;    APPLICANT:  Uehara, Hirotsugu
;    APPLICANT:  Eda, Soji
;    TITLE OF INVENTION:  Probe for Diagnosing Infectious Disease
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/920,828
;      FILING DATE:  29-AUG-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/362,577
;      FILING DATE:  27-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  19036/32420
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2362 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Staphylococcus epidermidis
;      STRAIN:  Clinical Isolate SE-3
US-08-920-828-5

  Query Match             8.1%;  Score 28.8;  DB 2;  Length 2362;
  Best Local Similarity   51.1%;  Pred. No. 12;
  Matches   93;  Conservative    0;  Mismatches   87;  Indels    2;  Gaps    1;

Qy      21 TCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATT 80
           ||   | ||  |  ||||    |   | || | |    || |||  | |||  |||||
Db    1074 TCATACAAAAGAAGATGTTGAGAAAATAGTAGATTCAAGTATTGACGGTTTACTTGTAGG 1133

Qy      81 GCTGGCGTTCATGCAGCAAAA--ACATAAGACACATTCTGGGGGTGATTACGGAGTCCCA 138
           |  ||| || ||| |   |||  || |||| ||  || | |    | |||  |      |
Db    1134 GGAGGCATTAATGAAAACAAATGACTTAAGTCAGTTTTTTGCCTAGTTTAAAGTTAAAGA 1193

Qy     139 ACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAAC 198
           | | | ||||     | || || | | |    || ||    |  |||| |||| |  |
Db    1194 AGAATCTCTATGATAGTTAAATTTTGTGGTTTTAAAACCGAAAGTGATATTAAGAAAATT 1253
```

Qy 199 AA 200
||
Db 1254 AA 1255

Search completed: March 11, 2003, 15:25:54
Job time : 72.5 secs

OM nucleic - nucleic search, using sw model

Run on: March 11, 2003, 13:59:58 ; Search time 743 Seconds
(without alignments)
320.925 Million cell updates/sec

Title: US-09-936-011-1
Perfect score: 357
Sequence: 1 atgtctagggaaataaccgc.....gtaatggtgcttgtggttaa 357

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 478924 seqs, 333959956 residues

Total number of hits satisfying chosen parameters: 957848

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	36.6	10.3	4039	10	US-09-925-301-371	Sequence 371, App
c	2	35	9.8	759	10	US-09-966-881-33	Sequence 33, Appl
c	3	32.4	9.1	2585	10	US-09-867-701-10881	Sequence 10881, A
c	4	32.2	9.0	234	10	US-09-864-761-28286	Sequence 28286, A
c	5	31.8	8.9	1128	9	US-09-938-842A-892	Sequence 892, App
c	6	31.8	8.9	2000	10	US-09-887-576-68	Sequence 68, Appl
	7	31.6	8.9	3124	10	US-09-925-302-340	Sequence 340, App
c	8	31.2	8.7	472	10	US-09-864-761-11710	Sequence 11710, A
c	9	31	8.7	493	10	US-09-864-761-5863	Sequence 5863, Ap
	10	31	8.7	730	10	US-09-764-864-248	Sequence 248, App
	11	31	8.7	753	10	US-09-764-853-177	Sequence 177, App
	12	31	8.7	818	10	US-09-764-864-676	Sequence 676, App
	13	31	8.7	873	10	US-09-764-853-178	Sequence 178, App
	14	30.6	8.6	263	10	US-09-923-876-4073	Sequence 4073, Ap
	15	30.4	8.5	2804	9	US-09-822-846-319	Sequence 319, App
c	16	30.4	8.5	127197	9	US-09-754-853A-1	Sequence 1, Appli
	17	30.4	8.5	155074	9	US-10-026-188-6	Sequence 6, Appli
c	18	30.2	8.5	574	10	US-09-864-761-228	Sequence 228, App
c	19	30.2	8.5	669	10	US-09-864-761-17051	Sequence 17051, A
c	20	30.2	8.5	2000	9	US-09-938-842A-3512	Sequence 3512, Ap
	21	30	8.4	401	10	US-09-864-761-3936	Sequence 3936, Ap
	22	30	8.4	446	10	US-09-864-761-20699	Sequence 20699, A
	23	30	8.4	926	10	US-09-925-301-287	Sequence 287, App
	24	29.6	8.3	287	9	US-09-535-459-1849	Sequence 1849, Ap
c	25	29.6	8.3	2361	10	US-09-815-242-8565	Sequence 8565, Ap
	26	29.6	8.3	4285	9	US-10-104-580-1	Sequence 1, Appli
c	27	29.6	8.3	9567	12	US-10-016-768-12	Sequence 12, Appl
c	28	29.6	8.3	9574	12	US-10-016-768-9	Sequence 9, Appli
	29	29.4	8.2	594	10	US-09-864-761-7908	Sequence 7908, Ap
c	30	29.4	8.2	1143	9	US-09-938-842A-1399	Sequence 1399, Ap
c	31	29.4	8.2	2043	9	US-09-938-842A-1337	Sequence 1337, Ap
c	32	29.4	8.2	10236	10	US-09-764-853-878	Sequence 878, App
	33	29.2	8.2	456	10	US-09-864-761-5185	Sequence 5185, Ap
	34	29.2	8.2	887	10	US-09-864-761-21484	Sequence 21484, A
	35	29.2	8.2	1075	10	US-09-864-761-19241	Sequence 19241, A
	36	29.2	8.2	1403	10	US-09-864-761-2513	Sequence 2513, Ap
c	37	29.2	8.2	1635	10	US-09-864-761-20241	Sequence 20241, A
c	38	29.2	8.2	1973	10	US-09-864-761-3471	Sequence 3471, Ap
	39	29.2	8.2	1981	10	US-09-864-761-4745	Sequence 4745, Ap
	40	29.2	8.2	24768	10	US-09-764-887-602	Sequence 602, App
	41	29	8.1	496	9	US-10-040-739-276	Sequence 276, App
	42	29	8.1	660	10	US-09-864-761-19488	Sequence 19488, A
	43	29	8.1	962	10	US-09-864-761-2772	Sequence 2772, Ap
	44	29	8.1	1370	10	US-09-902-684-12	Sequence 12, Appl
	45	29	8.1	1371	10	US-09-902-684-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-925-301-371
; Sequence 371, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 371
; LENGTH: 4039
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1085)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-371

Query Match 10.3%; Score 36.6; DB 10; Length 4039;
Best Local Similarity 49.7%; Pred. No. 0.36;
Matches 93; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

```
Qy     123 TGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGC 182
           |||   |   | ||||||    | |    ||  |  |  | || | |  | | ||  ||
Db     586 TGACCCCCTGGCACCAACACCCGCCAGTACTGTGACTTCCAAAGCCAGAGCCACATGTGC 645

Qy     183 TGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAG 242
           | ||   | | |  || ||    ||    |   || ||||  ||||||||     |
Db     646 TCATCAAACTTGCATTAAGCAGTTGGCGGGAGATGGCTGTGGAGCTGGGGGTTTAAGTGA 705

Qy     243 TAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCACT 302
           | ||    ||  ||  | ||| |   | ||    |  | ||| | |||| ||||   |
Db     706 TGGTTCTCTTTTGCTCCCTCTTYTGAGGGTAAAGCTACTGTCTTTCTTAAGAGTGTATTT 765
```

```
Qy     303 ATTACAA 309
           ||  |||
Db     766 ATGCCAA 772


RESULT 2
US-09-966-881-33/c
; Sequence 33, Application US/09966881
; Patent No. US20020120960A1
;   GENERAL INFORMATION:
;        APPLICANT: Seymour, Graham
;                   Bird, Colin
;                   Medina-Suarez, Rosybel
;        TITLE OF INVENTION: Genetic control Of Fruit Ripening
;        NUMBER OF SEQUENCES: 57
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Zeneca Ag Products Inc.
;             STREET: 1800 Concord Pike
;             CITY: Wilmington
;             STATE: DE
;             COUNTRY: USA
;             ZIP: 19850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/966,881
;             FILING DATE: 28-Sep-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/242,860
;             FILING DATE: 29-Mar-1999
;             APPLICATION NUMBER: GB 9618862.8
;             FILING DATE: 10-SEP-1996
;             APPLICATION NUMBER: GB 9708366.1
;             FILING DATE: 25-APR-1997
;             APPLICATION NUMBER: PCT/GB97/02424
;             FILING DATE: 08-SEP-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Hohenschutz, Liza D.
;             REGISTRATION NUMBER: 33,712
;             REFERENCE/DOCKET NUMBER: SEE 50183/UST
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (302) 886-1699
;   INFORMATION FOR SEQ ID NO: 33:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 759 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: unknown
;        MOLECULE TYPE: cDNA
;        IMMEDIATE SOURCE:
;             CLONE: U-U43
;        SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-09-966-881-33

  Query Match             9.8%;  Score 35;  DB 10;  Length 759;
  Best Local Similarity   46.9%;  Pred. No. 0.54;
  Matches  112;  Conservative    0;  Mismatches  126;  Indels    1;  Gaps    1;

Qy      53 GTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACAC 112
           | |||  | |  |   ||  |  | | |  |  | |   | |  |||   ||| ||
Db     741 GNGTTCNTTTATTTAATTCATGGGCAGTNTTTCCCTGGTTTCTTCAATTCCATNAGTGTT 682

Qy     113 ATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTA 172
           ||| |||||  |  | ||  ||||    ||| | |  ||  ||||   |  |    | |
Db     681 TTTCNGGGGGATAANATGGTTTCCCGGTTTTTCCNATNGGAAGTATGCNTGTAACTTGCA 622

Qy     173 CAAGATCAGCTG-ATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGG 231
           |||||||  | | ||   | |  | | | |        | |||| |  |   | ||  |
Db     621 NAAGATCATCCGGATCCCACTTTTGATAGTTTCAACTTTGACGGATTTGTCCGCTCCAGC 562

Qy     232 GGTAGCGTTAGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTT 290
           || | |||  | |||  | || |   | |  |  |       |  || |    ||  |||
Db     561 AGTGGNGTTTTTTGTCAATTTAGATCGGAGNTCNTNTACAGAAGTACTTNNACTTAGTT 503


RESULT 3
US-09-867-701-10881/c
; Sequence 10881, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
;  APPLICANT: Aglate, Paul A.
;  APPLICANT:  Jones, Robert
;  APPLICANT:  Harlocker, Susan L.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.497
;  CURRENT APPLICATION NUMBER: US/09/867,701
;  CURRENT FILING DATE:  2001-05-29
;  NUMBER OF SEQ ID NOS: 10912
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10881
;   LENGTH: 2585
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-867-701-10881

  Query Match             9.1%;  Score 32.4;  DB 10;  Length 2585;
  Best Local Similarity   50.0%;  Pred. No. 5.5;
  Matches   81;  Conservative    0;  Mismatches   81;  Indels    0;  Gaps    0;

Qy      42 TATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAA 101
           | || || ||  | ||| | || |   ||| ||| | ||  |    ||  |    |||||
Db    2578 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGCTTTTTTTTCAAAA 2519

Qy     102 ACATAAGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATA 161
           |||||   || |||       | ||||   |      |  ||  |     |  |||   |
Db    2518 ACATACTTCATATTTCCTCTTTTATTATATAAATATCAGTTTAACCTTTTGCTGTAAGAA 2459

Qy     162 TAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCA 203
           || | ||| |  ||||  | ||   ||   ||   | | |||
Db    2458 TATAAACGTTTTAAGAGGATCTTTGTTATTTATACAAATTCA 2417


RESULT 4
US-09-864-761-28286/c
; Sequence 28286, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
```

```
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28286
;   LENGTH: 234
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC004547.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.94
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.77
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.75
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.86
;   OTHER INFORMATION: NT HIT: Y11842.1, EVALUE 1.20e-01
;   OTHER INFORMATION: EST_HUMAN HIT: BE143078.1, EVALUE 3.60e-02
US-09-864-761-28286

  Query Match             9.0%;  Score 32.2;  DB 10;  Length 234;
  Best Local Similarity   46.9%;  Pred. No. 2.3;
  Matches  100;  Conservative    0;  Mismatches  113;  Indels    0;  Gaps    0;

Qy      27 CAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGC 86
           || | |   ||| | ||||       |||| ||     || |   |||||  |     |
Db     221 CACTTACTTTGTACTTATTTAGAACAGTGTCTGGTACATGACAAGCTTTGGCTATGATGA 162

Qy      87 GTTCATGCAGCAAAAACATAAGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTC 146
           | || || |  || |     ||  ||   ||  | | |         |    |
Db     161 TGAGTAGGAGAAAGATGATGATGGTGATAGTGATAGTAGTGAAGATAATGATAGTGGTGG 102

Qy     147 TAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCG 206
           | | ||| |||  |||   |||| |    |   | |||| || || || ||| | ||
Db     101 TGATGGTAGTAGTTATGATGACGATGGTGGTGATGATGATATTGATGGTGACAGTGATAA 42

Qy     207 TGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGT 239
           || |   ||  | ||||    ||| | | | ||
Db      41 TGGTGGTGGTAGTGGTGATGATGGTGATGGTGT 9


RESULT 5
US-09-938-842A-892/c
; Sequence 892, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
;  APPLICANT: Harper, Jeff
;  APPLICANT:  Kreps, Joel
;  APPLICANT:  Wang, Xun
;  APPLICANT:  Zhu, Tong
;  TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;  TITLE OF INVENTION:  SAME, AND METHODS OF USE
;  FILE REFERENCE: SCRIP1300-3
;  CURRENT APPLICATION NUMBER: US/09/938,842A
;  CURRENT FILING DATE:  2001-08-24
;  PRIOR APPLICATION NUMBER: US 60/227,866
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: US 60/264,647
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/300,111
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 892
;   LENGTH: 1128
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-892

  Query Match             8.9%;  Score 31.8;  DB 9;  Length 1128;
  Best Local Similarity   53.7%;  Pred. No. 5.9;
  Matches   66;  Conservative    0;  Mismatches   57;  Indels    0;  Gaps    0;

Qy      12 AATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTT 71
           || || | | | |   | |     |||   | |||||||||| ||||  |||||| | | |
Db     841 AAGAAGCTCGCCATGTATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTGT 782

Qy      72 CTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGGGGTGATTACGG 131
           | |    ||| ||  | |    ||| | | | | |  |||  ||  | |  |||  |
Db     781 CATGTGGTTGATGTGGATTTGACAGAAGAGAGAGAGCACAGTTTGAGTCGCCAATTCCCT 722

Qy     132 AGT 134
           |||
Db     721 TGT 719


RESULT 6
US-09-887-576-68/c
; Sequence 68, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
;  APPLICANT: Budworth, P.
;  APPLICANT:  Brown, D.
;  APPLICANT:  Chang, H.
;  APPLICANT:  Zhu, T.
;  APPLICANT:  Han, B.
;  APPLICANT:  Wang, X.
;  APPLICANT:  Cooper, Bret
;  TITLE OF INVENTION: Promoters for regulation of plant expression
;  FILE REFERENCE: 1360.001US1
;  CURRENT APPLICATION NUMBER: US/09/887,576
;  CURRENT FILING DATE:  2001-06-25
;  PRIOR APPLICATION NUMBER: US 60/213,848
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: US 60/214,087
;  PRIOR FILING DATE: 2000-06-23
;  PRIOR APPLICATION NUMBER: US 60/258,692
;  PRIOR FILING DATE: 2000-12-29
;  NUMBER OF SEQ ID NOS: 875
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 68
;   LENGTH: 2000
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-887-576-68

  Query Match             8.9%;  Score 31.8;  DB 10;  Length 2000;
  Best Local Similarity   53.7%;  Pred. No. 7.5;
  Matches   66;  Conservative    0;  Mismatches   57;  Indels    0;  Gaps    0;

Qy      12 AATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTT 71
           || || | | | |   | |     |||   | |||||||||| ||||  |||||| | | |
Db     522 AAGAAGCTCGCCATGTATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTGT 463

Qy      72 CTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGGGGTGATTACGG 131
           | |    ||| ||  | |    ||| | | | | |  |||  ||  | |  |||  |
Db     462 CATGTGGTTGATGTGGATTTGACAGAAGAGAGAGAGCACAGTTTGAGTCGCCAATTCCCT 403
```

```
Qy    132 AGT 134
          ||
Db    402 TGT 400


RESULT 7
US-09-925-302-340
; Sequence 340, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA104
;  CURRENT APPLICATION NUMBER: US/09/925,302
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05918
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 896
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 340
;   LENGTH: 3124
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-925-302-340

  Query Match             8.9%;  Score 31.6;  DB 10;  Length 3124;
  Best Local Similarity   51.4%;  Pred. No. 10;
  Matches   73;  Conservative    0;  Mismatches   69;  Indels    0;  Gaps    0;

Qy     23 GACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGC 82
          | |  ||   ||||||||  |  || |    || ||   |||||| ||  ||| ||||
Db   2609 GCCAAAACTGGAATGTGCAGAAAGTAGGCTTTGGTTAATTTGTGGATTCATTTTTATTTT 2668

Qy     83 TGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGGGGTGATTACGGAGTCCCAACAT 142
          || | ||  |  |      | | || |    |||||| |  ||||      | |    |
Db   2669 TGTCTTTGTTTAACTTTTTAAAAAATAAGATTTCTGGAGTAGATTGGTATATTCTGTTAA 2728

Qy    143 TTTCTAACGGTGGTATATATAG 164
             || || ||| |  || | |
Db   2729 AGACTTACAGTGATCCATTTTG 2750


RESULT 8
US-09-864-761-11710/c
; Sequence 11710, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 11710
;   LENGTH: 472
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC004547.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.94
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL =  1
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.77
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.75
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.86
US-09-864-761-11710

  Query Match             8.7%;  Score 31.2;  DB 10;  Length 472;
  Best Local Similarity   47.1%;  Pred. No. 6.2;
  Matches   96;  Conservative    0;  Mismatches  108;  Indels    0;  Gaps    0;

Qy     36 TGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCA 95
          ||| | ||||       |||| ||     || |   |||||  |     |       | |
Db    470 TGTACTTATTTAGAACAGTGTCTGGTACATGACAAGCTTTGGCTATGATGATGAGTAGGA 411

Qy     96 GCAAAAACATAAGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACGGTGG 155
          | || |  || |      ||  ||   ||  | | |         |    |  | | ||| |
Db    410 GAAAGATGATGATGGTGATAGTGATAGTAGTGAAGATAATGATAGTGGTGGTGATGGTAG 351

Qy    156 TATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCGTGCTTACGG 215
          ||  |||   |||| |    |    | |||| || || || ||| | ||  || |   ||
Db    350 TAGTTATGATGACGATGGTGGTGATGATGATATTGATGGTGACAGTGATAATGGTGGTGG 291

Qy    216 GTGCGGTGGGTCTGGGGGTAGCGT 239
             | ||||    ||| | | | ||
Db    290 TAGTGGTGATGATGGTGATGGTGT 267


RESULT 9
US-09-864-761-5863/c
; Sequence 5863, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
```

```
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 5863
;   LENGTH: 493
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC004169.3
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
US-09-864-761-5863

  Query Match             8.7%;  Score 31;  DB 10;  Length 493;
  Best Local Similarity   48.9%;  Pred. No. 7.3;
  Matches  113;  Conservative    0;  Mismatches  115;  Indels    3;  Gaps    1;

Qy      30 TAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTT 89
           | | ||||    |  || || ||||| | | || ||||   |  |  || |  ||| |
Db     424 TGACAATGGTGGTGGTGATGGTGGTGATAGGGTGGTGGTGATGGTGATAGTAGTGGTGGG 365

Qy      90 CATGCAGCAAAAACATAAGACACATTCTGGGGG---TGATTACGGAGTCCCAACATTTTC 146
             ||| |          | |      ||  ||| ||   || | | || |      |     |
Db     364 AATGGATGGTGGTGGTGATGGTGATGGTGGTGGTGTTGGTGATGGTGGTGATAGTGCTGG 305

Qy     147 TAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCG 206
           | | | ||||     | | || |||  | |    | || |  | || || | |||  | |
Db     304 TGAGGATGGTGATGGTGGTGATGGTGGAGGTGATGGTGGTAGTTATGGTGATAATGGTGG 245

Qy     207 TGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAGTAGTCGAGTTGGGCA 257
           |  |    |  | |||||   ||| | | | | |  ||||   | |||| |
Db     244 TAGTGGTGATGGTGGTGGTGTTGGTGATGGTGGTGATAGTATTGGTGGGAA 194
```

```
RESULT 10
US-09-764-864-248
; Sequence 248, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ23
;  CURRENT APPLICATION NUMBER: US/09/764,864
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 1792
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 248
;   LENGTH: 730
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-864-248

  Query Match             8.7%;  Score 31;  DB 10;  Length 730;
  Best Local Similarity   62.0%;  Pred. No. 8.6;
  Matches   49;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Qy      36 TGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCA 95
           ||| | || |||||| ||| ||||| |||| | ||| ||||  |||    | |    |
Db     535 TGTACATAATGTTGTGGGTCTTTGTTTTGTTGTTTTGTTTGCTTTGGGATCTTGTTGGAT 594

Qy      96 GCAAAAACATAAGACACAT 114
           |||   | ||| |  | ||
Db     595 GCACTTAGATATGGAAAAT 613


RESULT 11
US-09-764-853-177
; Sequence 177, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PJZ06
;  CURRENT APPLICATION NUMBER: US/09/764,853
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 939
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 177
;   LENGTH: 753
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-853-177

  Query Match             8.7%;  Score 31;  DB 10;  Length 753;
  Best Local Similarity   62.0%;  Pred. No. 8.7;
  Matches   49;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Qy      36 TGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCA 95
           ||| | || |||||| ||| ||||| |||| | ||| ||||  |||    | |    |
Db     593 TGTACATAATGTTGTGGGTCTTTGTTTTGTTGTTTTGTTTGCTTTGGGATCTTGTTGGAT 652

Qy      96 GCAAAAACATAAGACACAT 114
           |||   | ||| |  | ||
Db     653 GCACTTAGATATGGAAAAT 671


RESULT 12
US-09-764-864-676
; Sequence 676, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ23
;  CURRENT APPLICATION NUMBER: US/09/764,864
```

```
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 1792
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 676
;    LENGTH: 818
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-764-864-676

  Query Match              8.7%;  Score 31;  DB 10;  Length 818;
  Best Local Similarity   62.0%;  Pred. No. 9;
  Matches   49;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Qy      36 TGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCA 95
           ||| | || |||||| ||| ||||| |||| | ||| ||||  |||    | |    |
Db     653 TGTACATAATGTTGTGGGTCTTTGTTTTGTTGTTTTGTTTGCTTTGGGATCTTGTTGGAT 712

Qy      96 GCAAAAACATAAGACACAT 114
           |||   | ||| |  | ||
Db     713 GCACTTAGATATGGAAAAT 731


RESULT 13
US-09-764-853-178
; Sequence 178, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PJZ06
;  CURRENT APPLICATION NUMBER: US/09/764,853
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 939
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 178
;    LENGTH: 873
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-764-853-178

  Query Match              8.7%;  Score 31;  DB 10;  Length 873;
  Best Local Similarity   62.0%;  Pred. No. 9.2;
  Matches   49;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Qy      36 TGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCA 95
           ||| | || |||||| ||| ||||| |||| | ||| ||||  |||    | |    |
Db     712 TGTACATAATGTTGTGGGTCTTTGTTTTGTTGTTTTGTTTGCTTTGGGATCTTGTTGGAT 771

Qy      96 GCAAAAACATAAGACACAT 114
           |||   | ||| |  | ||
Db     772 GCACTTAGATATGGAAAAT 790


RESULT 14
US-09-923-876-4073
; Sequence 4073, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
;  APPLICANT: Lalgudi, Raghunath V.
;  APPLICANT:  Kamigaki, Laura Y. (Ito)
;  APPLICANT:  Sherman, Bradley K.
;  TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
;  FILE REFERENCE: PL-0012-1 CON
;  CURRENT APPLICATION NUMBER: US/09/923,876
;  CURRENT FILING DATE:  2001-08-06
;  PRIOR APPLICATION NUMBER: 09/298,329
;  PRIOR FILING DATE: 1999-04-21
;  PRIOR APPLICATION NUMBER: 60/085,331
;  PRIOR FILING DATE: 1998-05-05
;  NUMBER OF SEQ ID NOS: 6332
;  SOFTWARE: PERL Program
;
; SEQ ID NO 4073
;    LENGTH: 263
;    TYPE: DNA
;    ORGANISM: Zea mays
;    FEATURE:
;    NAME/KEY: misc_feature
;    OTHER INFORMATION: Incyte ID No. US20020013958A1 700454373H1
;    NAME/KEY: unsure
;    LOCATION: 71, 101, 261
;    OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-4073

  Query Match              8.6%;  Score 30.6;  DB 10;  Length 263;
  Best Local Similarity   52.8%;  Pred. No. 7.3;
  Matches   66;  Conservative    0;  Mismatches   59;  Indels    0;  Gaps    0;

Qy      38 TGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGC 97
           ||  | || |||  ||||||| | |    |  ||| | |       ||||||| |  |||
Db     122 TGAGTTTTTTTGGCGGTGTTTTTTTGTAAGGATTCATGGACGATTTGGCGTTTGTTTAGC 181

Qy      98 AAAAACATAAGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTA 157
           |||| | | |   |||||  | ||   | ||  ||||   ||       |   || |
Db     182 CTCAACAGATGTCTTGTTCTGATGATGGAAAAGGGATCCCGTGATAGCAACATTCTGAGA 241

Qy     158 TATAT 162
           ||||
Db     242 AATAT 246


RESULT 15
US-09-822-846-319
; Sequence 319, Application US/09822846
; Publication No. US20030027139A1
; GENERAL INFORMATION:
;  APPLICANT: Jacobs, Kenneth
;  APPLICANT:  McCoy, John M.
;  APPLICANT:  LaVallie, Edward R.
;  APPLICANT:  Collins-Racie, Lisa A.
;  APPLICANT:  Evans, Cheryl
;  APPLICANT:  Merberg, David
;  APPLICANT:  Treacy, Maurice
;  APPLICANT:  Agostino, Michael J.
;  APPLICANT:  Steininger II, Robert J.
;  APPLICANT:  Bowman, Michael R.
;  APPLICANT:  Spaulding, Vikki
;  APPLICANT:  Wong, Gordon G.
;  APPLICANT:  Clark, Hilary
;  APPLICANT:  Fechtel, Kim
;  APPLICANT:  Howes, Steven H.
;  APPLICANT:  Resnick, Richard J.
;  APPLICANT:  Gulukota, Kamalakar
;  APPLICANT:  Graham, James R.
;  APPLICANT:  Genetics Institute, Inc.
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
;  FILE REFERENCE: GIN 6400
;  CURRENT APPLICATION NUMBER: US/09/822,846
;  CURRENT FILING DATE:  2001-03-29
;  PRIOR APPLICATION NUMBER: 60/195,605
;  PRIOR FILING DATE: 2000-04-06
;  NUMBER OF SEQ ID NOS: 629
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 319
;    LENGTH: 2804
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-822-846-319

  Query Match              8.5%;  Score 30.4;  DB 9;  Length 2804;
  Best Local Similarity   61.2%;  Pred. No. 23;
  Matches   49;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy      34 AATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATG 93
           ||  | |||   ||| ||  |||||   || ||  ||| | | | ||| |  | || |||
```

```
Db    1089 AACATACCTTAAGTTTTTTTTGTTTTGTTTTTGTTTTTTTGTTTTTTGTTTTCCTTTATG 1148

Qy      94 CAGCAAAAACATAAGACACA 113
           || ||||| | || |  ||
Db    1149 AAGAAAAAATAAAATAGTCA 1168
```

Search completed: March 12, 2003, 08:11:06
Job time : 749 secs

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: March 12, 2003, 06:32:01 ; Search time 1191.5 Seconds
(without alignments)
4852.529 Million cell updates/sec

Title: US-09-936-011-1
Perfect score: 357
Sequence: 1 atgtctagggaaataaccgc.....gtaatggtgcttgtggttaa 357

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	42.4	11.9	1080	17	CNS00EPP	AL069494 Drosophil
	2	42.2	11.8	1101	17	CNS00LT2	AL078714 Drosophil
	3	38.6	10.8	1101	17	CNS0178M	AL107728 Drosophil
c	4	38	10.6	375	9	AJ388997	AJ388997 AJ388997
	5	38	10.6	1101	17	CNS0100X	AL098379 Drosophil
	6	37.2	10.4	367	17	AQ909830	AQ909830 GSSTc0830
	7	36.6	10.3	873	14	BQ215809	BQ215809 AGENCOURT
c	8	36.6	10.3	1007	17	CNS0079A	AL066675 Drosophil
	9	36.4	10.2	486	9	AA134045	AA134045 zo17d11.r
	10	36	10.1	328	10	BB224881	BB224881 BB224881
	11	36	10.1	1101	17	CNS00FOO	AL070854 Drosophil
	12	35.8	10.0	902	17	CNS006QP	AL065804 Drosophil
c	13	35.6	10.0	1207	17	CNS015ZN	AL106109 Drosophil
c	14	35.2	9.9	501	17	AQ447722	AQ447722 mgxb0010K
	15	35.2	9.9	529	13	BI743894	BI743894 kx42c08.y
	16	35.2	9.9	554	17	AQ523013	AQ523013 HS_5230_A
	17	35	9.8	884	17	CNS006U0	AL065923 Drosophil
	18	35	9.8	922	17	CNS0073W	AL066784 Drosophil
c	19	35	9.8	994	17	CNS00F9Q	AL070417 Drosophil
c	20	34.8	9.7	1101	17	CNS00LO0	AL068607 Drosophil
	21	34.6	9.7	477	14	BQ353390	BQ353390 PM1-HT045
	22	34.6	9.7	894	12	BG776437	BG776437 602663521
c	23	34.6	9.7	939	17	CNS00CNG	AL059400 Drosophil
	24	34.4	9.6	504	17	AQ529268	AQ529268 RPCI-11-3
c	25	34.4	9.6	552	17	AZ400908	AZ400908 1M0167G24
c	26	34.4	9.6	587	12	BE921791	BE921791 EST425560
	27	34.4	9.6	598	13	BJ032222	BJ032222 BJ032222
c	28	34.4	9.6	612	17	AG159078	AG159078 Pan trogl
c	29	34.4	9.6	645	17	AG000861	AG000861 Homo sapi
	30	34.4	9.6	660	14	BQ112425	BQ112425 EST598001
c	31	34.4	9.6	691	14	BQ112424	BQ112424 EST598000
c	32	34.4	9.6	720	12	BF863620	BF863620 963045E02
	33	34.4	9.6	721	13	BJ137716	BJ137716 BJ137716
	34	34.4	9.6	781	9	AA849835	AA849835 EST192602
c	35	34.4	9.6	1017	13	BI761612	BI761612 603046330
c	36	34.2	9.6	401	10	BB671302	BB671302 BB671302
	37	34.2	9.6	605	17	DR20K6S	AL739821 Danio rer
c	38	34.2	9.6	665	17	AZ863079	AZ863079 2M0171L01
c	39	34.2	9.6	910	17	CNS06JMH	AL401775 T7 end of
c	40	34	9.5	228	9	AU073923	AU073923 AU073923
c	41	34	9.5	254	9	AV240878	AV240878 AV240878
	42	34	9.5	325	9	AU271393	AU271393 AU271393
c	43	34	9.5	673	17	AG066509	AG066509 Pan trogl
c	44	34	9.5	1101	17	CNS006DU	AL065139 Drosophil
	45	33.8	9.5	418	17	B43757	B43757 HS-1058-A2-

ALIGNMENTS

RESULT 1
CNS00EPP
LOCUS CNS00EPP 1080 bp DNA linear GSS 04-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC: BACR29N07 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL069494
VERSION AL069494.1 GI:4949637
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1080)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial

EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES Location/Qualifiers
source 1. .1080
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR29N07"
/clone_lib="RPCI-98"
/note="end : TET3"

BASE COUNT 171 a 105 c 219 g 225 t 360 others
ORIGIN

Query Match 11.9%; Score 42.4; DB 17; Length 1080;
Best Local Similarity 20.7%; Pred. No. 1.2;
Matches 51; Conservative 102; Mismatches 93; Indels 0; Gaps 0;

```
Qy      45 TGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACA 104
           ::::::|::: ::   ::: ::: ::| :|  : :|: :|: :     : |: |  ::
Db     627 WDKWKKTKKKTKWGTRKWWTKKRGKDTRKTGTKTKTKKDGRKKGGKGRDTADGAKTDRKK 686

Qy     105 TAAGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAG 164
           |: |:   : :  |||:|:| ||  :|:|:   :| :|:|: :::   :|:  :  :::|
Db     687 TRKGDTKKRRDRAGGGKGDGGTTTDKGDGDWWADATDTWTKKKDDTTWWGDATDKRWWWG 746

Qy     165 AGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGG 224
           ::: :::: ::    |   :|:: |:|:|: :| :|: ::     |::  ::|: ::|:
Db     747 WRWAKDDDDWWAGRGAAGARADDRTDADARGWAARAKKDKTTAWATDWTTRKTRTDDTDT 806

Qy     225 GTCTGGGGGTAGCGTTAGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTC 284
           |:  :|:|: :| ::: | ||| :::  :|:  : || : : :    :  : :| |||
Db     807 GDGRKGRGKGWGKRKDGGGAGTRKRKGARGKGTDAAAAWAWGDTRTADKTADTKTTTTTA 866

Qy     285 TGTGTT 290
           | |:|:
Db     867 TTTKTK 872
```

RESULT 2
CNS00LT2
LOCUS CNS00LT2 1101 bp DNA linear GSS 14-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC: BACR48P19 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL078714
VERSION AL078714.1 GI:5102004
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (11-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES Location/Qualifiers
source 1. .1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR48P19"
/clone_lib="RPCI-98"
/note="end : TET3"

BASE COUNT 469 a 6 c 69 g 151 t 406 others
ORIGIN

Query Match 11.8%; Score 42.2; DB 17; Length 1101;
Best Local Similarity 13.7%; Pred. No. 1.3;
Matches 37; Conservative 127; Mismatches 107; Indels 0; Gaps 0;

```
Qy      28 AATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCG 87
           :: || ||  ::  :  :::::|:::| ::::::::::|| :::   ::: :   ::: :
Db     831 WDAAAAAAAAKKTKDKGKKKKKTKKKTTKKKKKKKKKKGGGKKDAAAKKKKKGTKKKKGK 890

Qy      88 TTCATGCAGCAAAAACATAAGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCT 147
           :   :|  :         ::: | |  :  ::::|: |  :  :| ::        ::|: |
Db     891 KGKGKGGGKGKGKGGKKDDDAAAAKKKGKKKKKGKGGGGKKKKGGKKGKGKGGKKKTKKT 950

Qy     148 AACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCGT 207
              ::::::|:|      ::   :| ||::| |  :: :::: :: :|     : |   |
Db     951 KKKKKKKKKADAAKGKTKKRAAADAAAADWTDAATKKKKKKKKDKKKTKTKKKDAAAAAT 1010

Qy     208 GCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAGTAGTCGAGTTGGGCAGCAACTTATT 267
           : ::  :::|: ::|:::: :::: : : :|| :::|: : |: :|:  :    |: ::
Db    1011 KKKKKKKKKTKKKKTKKKKKKKKKDADKKKKTTGKKDGKGKGGKGKGKKGKKKTKTKKKK 1070

Qy     268 GTGTTAGCTATTGTTTCTGTGTTAATAGTGT 298
           :::::   : || ::  ::::: |:: ::::
Db    1071 KKKKKKTKKKTTADKGKKKKKKAADDKKKKK 1101
```

RESULT 3
CNS0178M
LOCUS CNS0178M 1101 bp DNA linear GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence SP6 end of BAC BACN17P24 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL107728
VERSION AL107728.1 GI:5628032
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBeloBAC11.
FEATURES Location/Qualifiers
source 1. .1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACN17P24"
/clone_lib="DrosBAC"
/plasmid="pBeloBAC11"

```
                     /note="end : SP6"
BASE COUNT      275 a     87 c     95 g     82 t    562 others
ORIGIN

  Query Match             10.8%;  Score 38.6;  DB 17;  Length 1101;
  Best Local Similarity   18.4%;  Pred. No. 11;
  Matches   65;  Conservative  131;  Mismatches  158;  Indels    0;  Gaps    0;

Qy       2 TGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTG 61
          :::|:: | : : :|: | :|:    |: |:::: ||:  :    :|  :|  :  | |
Db     678 HBHCYMNGCDTWADABNGWDCKNAWNAWCAHVDRNATGBWVRVGCGBGSRVGSMBMNGCG 737

Qy      62 TTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGGG 121
           :: :  : :: |   :: :: : ::: : :  :::|:||:| :| | | |:: :| ::
Db     738 AGBBASNMBCSHGTYNCDKTHWTVTDWWAWAKRNBMRARAABAAWAWAAAAAWKWKGWRD 797

Qy     122 GTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAG 181
           :||  |   : ::   :  :|:|     : |  | | |:|: | : :  : : :  :
Db     798 VTGTSTSAARNDKKDDVTTYKTBTTNKKTKTTYTTKTTTMTWTAYDKDTGWRWGSKGYGT 857

Qy     182 CTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTA 241
           : : :| |  :|:::::     ::| |     ||: | :  ::|:  :|:|| |  |:
Db     858 YRBKKTGTNCDAVWVMYSYGNCKSGSGGSGSKGGKGGNRAGRVGDGGKGKGGGAAWGDST 917

Qy     242 GTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCAC 301
           :  ::: |    :  ::: :: :::| : :  : :|:|:|| :|:  ::  :  : :: :
Db     918 BCKKWBAAAAANRANHBSNRBNBMBTNKWKGBBWBAKTBTTNBTKNYYKTTBKYKYDSBB 977

Qy     302 TATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTT 355
           | :   ||  |   :: ::::: ::  || :   :|::    |::  :|: |::
Db     978 TKYNCAAAAAAAMNHVHBBBYBTSBKCAADWKTCKTBKKKGTGKKKGKGKTGKK 1031


RESULT 4
AJ388997/c
LOCUS       AJ388997                375 bp    mRNA    linear   EST 21-MAR-2000
DEFINITION  AJ388997 Medicago truncatula R108 Medicago truncatula cDNA clone
            MtNo372, mRNA sequence.
ACCESSION   AJ388997
VERSION     AJ388997.1  GI:6604264
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 375)
  AUTHORS   Gyorgyey,J., Vaubert,D., Jimenez-Zurdo,J.I., Charon,C., Troussard
            ,L., Kondorosi,A. and Kondorosi,E.
  TITLE     Analysis of Medicago truncatula nodule expressed sequence tags
  JOURNAL   Mol. Plant Microbe Interact. 13 (1), 62-71 (2000)
  MEDLINE   20120379
COMMENT     Contact: Gyorgyey J
            Institut des Sciences Vegetales
            Centre National de la Recherche Scientifique (CNRS)
            Av. de la Terrasse Bat.23, Gif-sur-Yvette, 91198, FRANCE
            Additional information about these EST clones can be downloaded
            from
            http://www.cnrs-gif.fr/isv/AK/index.html. (URL provided by the
            author).
FEATURES             Location/Qualifiers
     source          1. .375
                     /organism="Medicago truncatula"
                     /strain="R108"
                     /db_xref="taxon:3880"
                     /clone="MtNo372"
                     /clone_lib="Medicago truncatula R108"
                     /tissue_type="symbiotic root nodule"
                     /dev_stage="developing young nodule"
                     /note="Vector: NOLLY Lambda HybriZAP"
BASE COUNT      146 a     44 c     46 g    138 t      1 others
ORIGIN

  Query Match             10.6%;  Score 38;  DB 9;  Length 375;
  Best Local Similarity   52.2%;  Pred. No. 16;
  Matches   83;  Conservative    0;  Mismatches   76;  Indels    0;  Gaps    0;

Qy      42 TATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAA 101
           | || || ||  | |||  | || |   ||| ||| | ||  |    |  | |||  ||||
Db     374 TTTTTTTTTTTTTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAAGGCAAGAAAA 315

Qy     102 ACATAAGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATA 161
           | ||||  || | | || ||   ||||   ||| | |  |  || ||    |  || | |
Db     314 ATATAAATCAAACTGTGAGGAAAATTATTAAGTACAATTAAATTGTATACTTATTACAAA 255

Qy     162 TAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAA 200
           |    |   |  || |  ||   | ||    || || ||
Db     254 TTATTATTTTTTAATAAAAGGAAACTTGCCAAGAAAGAA 216


RESULT 5
CNS0100X
LOCUS       CNS0100X                1101 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN03G04 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL098379
VERSION     AL098379.1  GI:5609990
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACN03G04"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : SP6"
BASE COUNT      195 a    108 c    131 g    161 t    506 others
ORIGIN

  Query Match             10.6%;  Score 38;  DB 17;  Length 1101;
  Best Local Similarity   17.7%;  Pred. No. 16;
  Matches   56;  Conservative  131;  Mismatches  129;  Indels    0;  Gaps    0;

Qy      42 TATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAA 101
           |:|::: :  | : ::::|: ::::  |:| :|:    |: : : |    :    :::|
Db     506 TWTDVWADNAGWWTDKDKTKAKKKRTKTDTKKTKTRTGTKKKTDRGMRKTDTWGWGDWRA 565

Qy     102 ACATAAGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATA 161
           :  : |:| : |   ::: :|: :|:| : :|:      :||:: :|:  :::::::: :|
Db     566 DRTDGARADRAANAKDKDADGKADTDATKTDGWGTTTKADTTDDTDADWWDWKDWDWKDA 625

Qy     162 TAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGG 221
           : :||: :|:    :    : |: :::: :: :|     : :: ::  ::: |:: : ::
Db     626 DTKAGRKRGWGWGKKGGWKKTTKTKKKKTDKTKTGTDTKWGWKDKWTDKKRDGKRGDGDR 685
```

```
Qy     222 TGGGTCTGGGGGTAGCGTTAGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGT 281
           :::|| :|::: | : :| |: :| : | || ||: : || : | || |
Db     686 KDKGTGGDGDSKAKGDRAWAWRAKATRAAAADATAAATAAKATAWATKTTTKKTTTTTTT 745

Qy     282 TTCTGTGTTAATAGTGTCACTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAA 341
           :| | | || | :| | : : :: : ::: : :|: : : ::| : :::|:::
Db     746 KTTTTTTTTTTTKKTTTARKADRDDDDKAKRDGDKDGDWDDADKDWKAGRDRDDDTRKWG 805

Qy     342 TGGTGCTTGTGGTTAA 357
           ::: :::: :::::
Db     806 WKDRAGKKKKAKKKDH 821


RESULT 6
AQ909830
LOCUS       AQ909830                367 bp    DNA     linear   GSS 09-JAN-2001
DEFINITION  GSSTc08309 Trypanosoma cruzi random genomic library Trypanosoma
            cruzi genomic clone G37J17, DNA sequence.
ACCESSION   AQ909830
VERSION     AQ909830.3  GI:10135556
KEYWORDS    GSS.
SOURCE      Trypanosoma cruzi.
  ORGANISM  Trypanosoma cruzi
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma; Schizotrypanum.
REFERENCE   1  (bases 1 to 367)
  AUTHORS   Aguero,F., Verdun,R., Frasch,A.C.C. and Sanchez,D.O.
  TITLE     A random sequencing approach for the analysis of the trypanosoma
            cruzi genome: general structure, large gene and repetitive DNA
            families, and gene discovery
  JOURNAL   Genome Res. 10 (12), 1996-2005 (2000)
  MEDLINE   20568489
COMMENT     On Sep 14, 2000 this sequence version replaced gi:9374600.
            Contact: Sanchez D.O.
            Instituto de Investigaciones Biotecnologicas (Univ. Nac. de Gral
            San Martin)
            Av. Gral Paz entre Albarellos y Constituyentes, INTI edificio 24
            CP(1650) San Martin, Prov. de BS AS. Argentina
            Tel: 54-11-4580-7255 ext 309
            Fax: 54-11-4752-9639
            Email: dsanchez@iib.unsam.edu.ar
            Sequences were basecalled with phred and vector was masked with
            crossmatch (see http://genome.washington.edu). Sequences were then
            trimmed from both ends to remove low quality bases and masked
            vector.
            Seq primer: T7
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .367
                     /organism="Trypanosoma cruzi"
                     /strain="CL-Brener"
                     /db_xref="taxon:5693"
                     /clone="G37J17"
                     /clone_lib="Trypanosoma cruzi random genomic library"
                     /cell_type="epimastigote"
                     /note="Vector: pBS(-) (Stratagene); T. cruzi DNA was
                     randomly sheared using a nebulizer and the 1 to 2 Kb range
                     was gel purified and cloned into the dephosphoryated
                     HincII site of the vector"
BASE COUNT       78 a     56 c     83 g    150 t
ORIGIN

  Query Match             10.4%;  Score 37.2;  DB 17;  Length 367;
  Best Local Similarity   54.3%;  Pred. No. 25;
  Matches   75;  Conservative    0;  Mismatches   63;  Indels    0;  Gaps    0;

Qy      33 GAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCAT 92
           | | || | | || || ||||||| |||| || || ||||| || | | ||
Db     114 GCAGGTACTTTTTTTTTTTGGTGTGCGTGTTGTTGTTGTTTGTTTTTGTTGTTTTTGA 173

Qy      93 GCAGCAAAAACATAAGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACGG 152
             |||| | | ||| | | | || || || | ||| |||| | |||
Db     174 TACTCAAACAAACAAGCAAAAAAAAGAAGGAGAAAACAAATTCCTTTTTTTTTTTTTCGG 233
```

```
Qy     153 TGGTATATATAGAGACGG 170
             | |  |||| ||
Db     234 GTGGAGGAGGAGAGAGGG 251


RESULT 7
BQ215809
LOCUS       BQ215809                873 bp    mRNA    linear   EST 02-MAY-2002
DEFINITION  AGENCOURT_7574890 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6059837
            5', mRNA sequence.
ACCESSION   BQ215809
VERSION     BQ215809.1  GI:20397219
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 873)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13327  row: c  column: 06
            High quality sequence stop: 645.
FEATURES             Location/Qualifiers
     source          1. .873
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6059837"
                     /clone_lib="NIH_MGC_68"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.8 kb.   Library constructed by Life
                     Technologies. "
BASE COUNT      203 a    224 c    162 g    284 t
ORIGIN

  Query Match             10.3%;  Score 36.6;  DB 14;  Length 873;
  Best Local Similarity   49.7%;  Pred. No. 36;
  Matches   93;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy     123 TGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGC 182
           |||   |   | |||||| | |   || | | | || | | | | || ||
Db      33 TGACCCCCTGGCACCAACACCCGCCAGTACTGTGACTTCCAAAGCCAGAGCCACATGTGC 92

Qy     183 TGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAG 242
           | ||  | | | |  || ||   ||   |  || |||| |||||||    |
Db      93 TCATCAAACTTGCATTAAGCAGTTGGCGGGAGATGGCTGTGGAGCTGGGGGTTTAAGTGA 152

Qy     243 TAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCACT 302
           | ||   ||  ||  |  ||| |  | ||    |  | ||| | |||| |||||   |
Db     153 TGGTTCTCTTTTGCTCCCTCTTTTGAGGGTAAAGCTACTGTCTTTCTTAAGAGTGTATTT 212

Qy     303 ATTACAA 309
           ||  |||
Db     213 ATGCCAA 219


RESULT 8
CNS0079A/c
LOCUS       CNS0079A                1007 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC #
```

```
            BACR15M01 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL066675
VERSION     AL066675.1  GI:4945441
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1007)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .1007
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACR15M01"
                     /clone_lib="RPCI-98"
                     /note="end : T7"
BASE COUNT      314 a    243 c     70 g    197 t    183 others
ORIGIN

  Query Match           10.3%;  Score 36.6;  DB 17;  Length 1007;
  Best Local Similarity  35.3%;  Pred. No. 36;
  Matches   73;  Conservative   39;  Mismatches   95;  Indels    0;  Gaps    0;

Qy      46 GTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACAT 105
           |: ||| :::|| : ||:|:| || |:|| ||:  |:  |:| |   :: ||||| |
Db     951 GKGGTTTKKKTTGTKTTTKTKGGGTTTTKTGGATKTTTKTTGKTTAAAAWRAAAAAAAAA 892

Qy     106 AAGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGA 165
           || | | |    :   |  | | ||| |      ::   |  ||  ::|  :|:||:
Db     891 AAAAAAAAARAGRAATGGAAAAAGGGAATTRATGTTGKKGGAAGGGGGTDDAAGWAKAGH 832

Qy     166 GACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGG 225
             :   ::: |::| ::     |: :|| | : :  |  | |   | | : | | :|||
Db     831 TGMYCGCMHHMAWMACMWMYMAATWCHAGGAYVGWGAAATGAGAGMAGGMSMGGGAWGGG 772

Qy     226 TCTGGGGGTAGCGTTAGTAGTCGAGTT 252
               ||   |||::: |||:|||
Db     771 AKCTTTTTTAWTTTTAKKWTTCGWGTT 745


RESULT 9
AA134045
LOCUS       AA134045                 486 bp    mRNA    linear   EST 27-NOV-1996
DEFINITION  zo17d11.r1 Stratagene colon (#937204) Homo sapiens cDNA clone
            IMAGE:587157 5', mRNA sequence.
ACCESSION   AA134045
VERSION     AA134045.1  GI:1691275
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 486)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 318.
FEATURES             Location/Qualifiers
     source          1. .486
                     /organism="Homo sapiens"
                     /db_xref="GDB:4619459"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:587157"
                     /clone_lib="Stratagene colon (#937204)"
                     /tissue_type="tumor"
                     /cell_line="T84 carcinoma cell line"
                     /lab_host="SOLR cells (kanamycin resistant)"
                     /note="Organ: colon; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. T-84 colonic epithelial cell line.  Average
                     insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor
                     sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT      136 a    119 c     89 g    138 t      4 others
ORIGIN

  Query Match           10.2%;  Score 36.4;  DB 9;  Length 486;
  Best Local Similarity  50.6%;  Pred. No. 40;
  Matches   88;  Conservative    0;  Mismatches   86;  Indels    0;  Gaps    0;

Qy     136 CCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGT 195
           ||||||    | |    ||  |  |  | || | |  | | ||  ||| ||   | | |
Db     119 CCAACACCCGCCAGTACTGTGACTTCCAAAGCCAGAGCCACATGTGCTCATCAAACTTGC 178

Qy     196 AACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAGTAGTCGAGTTGGG 255
           |  || ||    ||    |   || ||||  ||||||||     |  | ||    ||  |
Db     179 ATTAAGCAGTTGGCGGGAGATGGCTGTGGAGCTGGGGGTTTAAGTGATGGTTCTCTTTTG 238

Qy     256 CAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCACTATTACAA 309
           |  |  ||| |   | ||    |  | ||| | |||| |||||   |||  |||
Db     239 CTCCCTCTTTTGAGGGTAAAGCTACTGTCTTTCTTAAGAGTGTATTTATGCCAA 292


RESULT 10
BB224881
LOCUS       BB224881                 328 bp    mRNA    linear   EST 01-JUL-2000
DEFINITION  BB224881 RIKEN full-length enriched, adult male aorta and vein Mus
            musculus cDNA clone A530088K07 3' similar to D43951 Homo sapiens
            mRNA for KIAA0099 protein, mRNA sequence.
ACCESSION   BB224881
VERSION     BB224881.1  GI:8893493
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 328)
  AUTHORS   Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci
            ,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
```

Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

TITLE RIKEN Mouse ESTs (Konno,H., et al.)

JOURNAL Unpublished (2000)

COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

FEATURES Location/Qualifiers
source 1. .328
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="A530088K07"
/clone_lib="RIKEN full-length enriched, adult male aorta and vein"
/sex="male"
/tissue_type="aorta and vein"
/dev_stage="adult"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT 83 a 70 c 50 g 125 t
ORIGIN

Query Match 10.1%; Score 36; DB 10; Length 328;
Best Local Similarity 51.9%; Pred. No. 51;
Matches 81; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

```
Qy    177 ATCAGCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAG 236
          ||||  |  ||| |  | |||    | || || | || | ||    ||  || |
Db     82 ATCATTTTTTTTAATGACTAATTCCCCTCCTGTTCACTGCTGTTTAGGTACTTGTACATA 141

Qy    237 CGTTAGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGT 296
          ||||  |||  | |||||  || ||||  | ||||   ||||| ||| | ||     |
Db    142 GGTTATCAGTGTACTTGGGTGGCCACTTGCTTTGTTCTATATTGTATCTTTATTCCACAT 201

Qy    297 GTCACTATTACAACGATTAAGGTCTCCACCAGAACA 332
          ||    | |   | | |    |  |  || | || |
Db    202 GTGGACACTTTCAGGGTGGTTGGTTTAACAAAAAAA 237
```

RESULT 11
CNS00FOO
LOCUS CNS00FOO 1101 bp DNA linear GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC: BACR31F03 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL070854
VERSION AL070854.1 GI:4950896
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES Location/Qualifiers
source 1. .1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR31F03"
/clone_lib="RPCI-98"
/note="end : T7"

BASE COUNT 167 a 84 c 86 g 155 t 609 others
ORIGIN

Query Match 10.1%; Score 36; DB 17; Length 1101;
Best Local Similarity 14.6%; Pred. No. 51;
Matches 43; Conservative 127; Mismatches 124; Indels 0; Gaps 0;

```
Qy     22 CGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTG 81
          :| ::::  ::  ::|::    |:::  ::::  |: :   ||:  :   :||:| :|:
Db    804 MGGMMMMGAMMTGVKGKKGTGGTKVBGTGYKKKNGTKNKMNMGTKWTDMNTKTTKTTKTK 863

Qy     82 CTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGGGGTGATTACGGAGTCCCAACA 141
           ::   |:  |:      :    : :|   :::  : : :: | :  ::|:    |: |
Db    864 TKKTTATWTKTKTGMKKKDKKKKKKTKAKKKMKKKAKAKAKKMAGDGAKDAKATAAAWAA 923

Qy    142 TTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAAT 201
          : :   |:       |:| |:| ::  : :   |::|  : ::|  : :|:|:::|  ::
Db    924 KAKAKAADAAAAWAAADAAAKAADKGTRWDGAKAKRTKKDAKKARAKADAKADDRAKKDK 983

Qy    202 CATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAGTAGTCGAGTTGGGCAGCAA 261
```

```
                : ::: ::  ::: : :::|::: ::::| :|: :::| : :: :|:::: :     :
Db    984 WKKKKKKKKKKRDKKGKKDKKGKDKDKKKKGAKAKAKKKATKKKKDKAKKDKAKDKAKKR 1043

Qy    262 CTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCACTATTACAACGATTA 315
          ::  : : :|:: |: : ::: : :::|: : :::|   :||:| |  :: ||
Db   1044 KDKKAKWKTKTWDATDADTDKKAKAKDKTWKDKKWKTATAKATDATAKAKWATA 1097


RESULT 12
CNS006QP
LOCUS       CNS006QP                 902 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR14J09 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL065804
VERSION     AL065804.1  GI:4944772
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 902)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .902
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACR14J09"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
BASE COUNT      298 a     50 c     61 g    142 t    351 others
ORIGIN

  Query Match           10.0%;  Score 35.8;  DB 17;  Length 902;
  Best Local Similarity   21.2%;  Pred. No. 57;
  Matches   54;  Conservative   93;  Mismatches  108;  Indels    0;  Gaps    0;

Qy     36 TGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCA 95
          |:::  : :: :||:|:::| :||:: || ::  |:: :|:::| |:| ::| |::  ::
Db    524 TKKKTKKKKKTKTGKTKKTTKTTKKTTTTKKTKTKKTKTKKTKTKGKKKGKGKKKTBKGG 583

Qy     96 GCAAAAACATAAGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACGGTGG 155
           :       :  |     :  :|||||   ::  :  :|       ::|: :    :::
Db    584 TBGGGGKGKBGGGGGGGGKGGKGGGGGGTTKKKTKTTKTKKTKTTTKKTKTKKTKTTKKK 643

Qy    156 TATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCGTGCTTACGG 215
          :   | : : |   :: :    :    |: :::|  |  |     :  : |||::| |:|
Db    644 KKGTTKYTKKGBKTKKTBTCCKKKTTTTBKKBKTTTTKTTKTKTKYKTKTGTGBBTYCSG 703

Qy    216 GTGCGGTGGGTCTGGGGGTAGCGTTAGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGC 275
          :|:::::::::: ::|||:  | ||: :| ::   :|::: : ::   :: ||:| ||
Db    704 KTKBBKKKKKKGKBGGGKGKGKGTBKBTKKKKYTBTBBKYBTBBYKTKYBTTKTTTTBTT 763

Qy    276 TATTGTTTCTGTGTT 290
          : ||::::| |:: ::
Db    764 BCTTBYBTTTKYTYB 778


RESULT 13
CNS015ZN/c
LOCUS       CNS015ZN                 1207 bp    DNA     linear   GSS 26-JUL-1999
DEFINITION  Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN15A19 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL106109
VERSION     AL106109.1  GI:5620099
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1207)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBeloBAC11.
FEATURES             Location/Qualifiers
     source          1. .1207
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACN15A19"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"
                     /note="end : SP6"
BASE COUNT      274 a     257 c     87 g    281 t    308 others
ORIGIN

  Query Match           10.0%;  Score 35.6;  DB 17;  Length 1207;
  Best Local Similarity   31.4%;  Pred. No. 65;
  Matches   59;  Conservative   47;  Mismatches   82;  Indels    0;  Gaps    0;

Qy    118 GGGGGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGA 177
          |||| :: ::  || :      ||||| | | :: |::: |:| | ::|  : | :::
Db    847 GGGGTKKKKKKGGGGRGGKTTTAATTTTTTTAARRRGKWRAAWAAAWDRARARAAARDKG 788

Qy    178 TCAGCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGC 237
            :| | |    :|:|  || ||| || ::  |::    ::  :::::    ::|:|
Db    787 ATVGDTWAAAAARAWAAAAAAAATGATRRWTKTWWTTTWKDTTTWDRRWRRAAKKGWAAK 728

Qy    238 GTTAGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTG 297
          ||| :|  | | |  :||  : |   | ||: ::|:  |  | :|| : | | :   :||
Db    727 GTTTKTTTTKGGGGGKGGGGKWAGGATTTTKARWTRTTTGGTTKTTTWTTTTGRKGGKTG 668

Qy    298 TCACTATT 305
          :   |:||
Db    667 KGKTTDTT 660


RESULT 14
AQ447722/c
LOCUS       AQ447722                 501 bp    DNA     linear   GSS 08-APR-1999
DEFINITION  mgxb0010K03f CUGI Rice Blast BAC Library Magnaporthe grisea genomic
            clone mgxb0010K03f, DNA sequence.
ACCESSION   AQ447722
VERSION     AQ447722.1  GI:4576859
KEYWORDS    GSS.
```

```
SOURCE      Magnaporthe grisea.
  ORGANISM  Magnaporthe grisea
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
            Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
REFERENCE   1  (bases 1 to 501)
  AUTHORS   Yu,Y., Zhu,H., Boyd,C.A., Gaudette,B., Gayle,A., Kingsbury,R.,
            Phillips,K., Sasinowski,M, Wing,R.A. and Dean,R.A.
  TITLE     A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
            Genome
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Dean RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson Universiy, Clemson, SC 29634
            Tel: 864 656 5737
            Fax: 864 656 4293
            Email: rdean@clemson.edu
            Seq primer: TAATACGACTCACTATAGGG
            Class: BAC ends
            High quality sequence stop: 357.
FEATURES             Location/Qualifiers
     source          1. .501
                     /organism="Magnaporthe grisea"
                     /strain="70-15"
                     /db_xref="taxon:148305"
                     /clone="mgxb0010K03f"
                     /clone_lib="CUGI Rice Blast BAC Library"
                     /tissue_type="Protoplasts"
                     /lab_host="E. coli DH10B"
                     /note="Vector: pBACWICH; Site_1: HindIII; Site_2: HindIII;
                     Rice blast is one of the most devestating fungal diseases
                     of rice  world wide.  It is a filamentous ascomycete with
                     a haploid genome (n=7)  of approximately 40 Mbp.  Rice
                     blast is an important model fungal pathogen  for studying
                     numerous aspects of the fungal-host interaction.  In
                     order to facilitate genome wide analysis, a BAC library
                     containing  9216 clones with an average insert size of 130
                     kbp was constructed.  This library represents greater
                     than 25X genome coverage.  High density  colony filters
                     are available upon request."
BASE COUNT      173 a    107 c    103 g    118 t
ORIGIN
```

```
  Query Match             9.9%;  Score 35.2;  DB 17;  Length 501;
  Best Local Similarity   52.0%;  Pred. No. 82;
  Matches   79;  Conservative    0;  Mismatches   73;  Indels    0;  Gaps    0;

Qy      25 CCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTG 84
           || || ||    ||   | |  ||||| ||||| |    |||  | | |   ||| |
Db     185 CCGAACAAATGGGTTTTTTTATTTGTTTTTGTTTTGGAATCGGCGATTTATCGTTTGGTT 126

Qy      85 GCGTTCATGCAGCAAAAACATAAGACACATTCTGGGGGTGATTACGGAGTCCCAACATTT 144
             |  | | | | |   ||    | | |||  | |  ||| | |  ||  | ||||  |
Db     125 TCTGTAACGTATCGGCAAACCCATAAACAAACAGAAGGTAAAAATAGACCCTCAACGGTC 66

Qy     145 TCTAACGGTGGTATATATAGAGACGGTACAAG 176
           |||  || ||    ||| | |||||   ||||
Db      65 TCTGGCGTTGTCTCATAAATAGACGTGTCAAG 34
```

```
RESULT 15
BI743894
LOCUS       BI743894                529 bp    mRNA    linear   EST 25-SEP-2001
DEFINITION  kx42c08.y1 Parastrongyloides trichosuri FL pAMP1 v1 Chiapelli
            McCarter Parastrongyloides trichosuri cDNA 5' similar to TR:Q9VRD9
            Q9VRD9 CG1753 PROTEIN. ;, mRNA sequence.
ACCESSION   BI743894
VERSION     BI743894.1  GI:15765696
KEYWORDS    EST.
SOURCE      Parastrongyloides trichosuri.
  ORGANISM  Parastrongyloides trichosuri
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Panagrolaimoidea; Strongyloididae; Parastrongyloides.
REFERENCE   1  (bases 1 to 529)
  AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T.
            , Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.
            , Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R.
            , Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe
            ,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.
            , Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
            Wilson,R.
  TITLE     The Washington Univ. Nematode EST Project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            The library was constructed by Brandi Chiapelli and Dr. James
            McCarter (bchiapel@watson.wustl.edu & jmccarte@watson.wustl.edu) at
            Washington University, St. Louis. DNA Sequencing by: Washington
            University Genome Sequencing Center St. Louis.
            Seq primer: -40RP from Gibco
            High quality sequence stop: 395.
FEATURES             Location/Qualifiers
     source          1. .529
                     /organism="Parastrongyloides trichosuri"
                     /db_xref="taxon:131310"
                     /clone_lib="Parastrongyloides trichosuri FL pAMP1 v1
                     Chiapelli McCarter"
                     /dev_stage="Free Living"
                     /lab_host="DH10B"
                     /note="Vector: pAMP1 (Gibco); Site_1: NotI; Site_2: SalI;
                     The library was constructed by Brandi Chiapelli and Dr.
                     James  McCarter at Washington University, St. Louis. The
                     cDNA was  made by using Dynabead oligo-dT priming (Dynal).
                     PCR based  library using a modified protocol from the
                     SMART PCR cDNA Synthesis  Kit from Clontech. Directionally
                     cloned into the UDG sites of pAMP1.  Nematodes were
                     provided by Dr. Warwick Grant of AgRearch, New Zealand
                     (warwick.grant@agresearch.co.nz)."
BASE COUNT      201 a     67 c    107 g    154 t
ORIGIN
```

```
  Query Match             9.9%;  Score 35.2;  DB 13;  Length 529;
  Best Local Similarity   48.5%;  Pred. No. 82;
  Matches   97;  Conservative    0;  Mismatches  103;  Indels    0;  Gaps    0;

Qy      70 TTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGGGGTGATTAC 129
           || ||  |  ||| ||   |   | | ||  ||||  |     |      | | |
Db     115 TTGTTATTGGTGCAGGAACTGGAGGAACAGTAACAGGAATTGGAAAAAAAAATTAAAGAA 174

Qy     130 GGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTT 189
           |||| ||||||| |   |  | ||||| | ||  ||| ||| ||  || |||| |   |
Db     175 AGAGTTCCAACATGTCAAATTGTTGGTACAGATCCAGAAGGTTCATTATTAGCTAACCCT 234

Qy     190 AATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAGTAGTCGA 249
           ||| |||  |     | | ||||     |  ||  | ||   |    ||  |     |
Db     235 GATATTAAAGAAATCTCTTCATACGAAGTTGAAGGTACAGGATATGATTTTGTTCCAGCA 294

Qy     250 GTTGGGCAGCAACTTATTGT 269
           |||    |   |  ||||||
Db     295 GTTCTTGACAGAAGTATTGT 314
```

```
Search completed: March 12, 2003, 10:11:16
Job time : 1199.5 secs
```

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: March 11, 2003, 15:26:24 ; Search time 1159 Seconds
(without alignments)
8964.367 Million cell updates/sec

Title: US-09-936-011-3
Perfect score: 357
Sequence: 1 atgtctagggaaataaccgc.....gtaatggtgcttgtggttaa 357

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		357	100.0	357	6	AX039060	AX039060 Sequence
2		347.4	97.3	357	6	AX039058	AX039058 Sequence
3		347.4	97.3	4612	14	TOBNYVV2	X04197 Beet necrot
4		345.8	96.9	1922	14	BNYVMTP1	X75575 Beet necrot
5		328.2	91.9	739	6	E59890	E59890 Rhizomania-
6		328.2	91.9	4544	14	AF197556	AF197556 Beet necr
7		328.2	91.9	4609	6	E59889	E59889 Rhizomania-
8		328.2	91.9	4609	14	BNYKPRO2	D84411 Beet necrot
9		323.4	90.6	4544	14	AF197547	AF197547 Beet necr
10		321.8	90.1	1922	14	BNYVMTP2	X75574 Beet necrot
11		315.4	88.3	739	6	E59891	E59891 Rhizomania-
12		208.2	58.3	4616	14	AF061869	AF061869 Beet soil
13		65.4	18.3	3065	14	LRVBABCD	Z46351 Lychnis rin
14		59.6	16.7	3005	14	BSBV3GNS	Z66493 Beet soil-b
15		53	14.8	3591	14	PSZBETACD	M81486 Poa semilat
16		50.4	14.1	2529	14	BVQ223598	AJ223598 Beet viru
17		48.6	13.6	2417	14	D86638	D86638 Broad bean
18		47.4	13.3	2962	14	PXMTGBP	D30753 Potato mop-
19		47.4	13.3	2964	14	PMO277556	AJ277556 Potato mo
20		46.2	12.9	3215	14	BSU35769	U35769 Barley stri
21		46.2	12.9	3237	14	BSU35772	U35772 Barley stri
22		46.2	12.9	3289	14	HOBMSVRB	X03854 Barley stri
23		46	12.9	4290	14	AF239729	AF239729 Indian pe
24		45.8	12.8	4504	14	PNWRNAII	L07269 Peanut clum
25	c	45.6	12.8	204615	2	AC112254	AC112254 Homo sapi
26		44.6	12.5	3239	14	BSU35770	U35770 Barley stri
27		44.6	12.5	3249	14	BSU35771	U35771 Barley stri
28	c	43.2	12.1	32070	2	AC116983	AC116983 Dictyoste
29		42.4	11.9	2325	14	NVMV2	D00906 Nicotiana v
30	c	42	11.8	164140	2	AC106093	AC106093 Rattus no
31	c	41	11.5	171438	2	AC099231	AC099231 Rattus no
32	c	41	11.5	179846	2	AC119387	AC119387 Rattus no
33	c	39.6	11.1	136587	9	AL445309	AL445309 Human DNA
34		39.6	11.1	175495	2	AC023938	AC023938 Mus muscu
35	c	39.6	11.1	178870	2	AC021555	AC021555 Homo sapi
36	c	39.6	11.1	215738	2	AC113274	AC113274 Mus muscu
37		39.6	11.1	258120	2	AC114408	AC114408 Mus muscu
38		38.8	10.9	128433	2	AC124151	AC124151 Oryza sat
39	c	38.4	10.8	171031	2	AC104020	AC104020 Homo sapi
40	c	38.4	10.8	202504	2	AC127585	AC127585 Mus muscu
41	c	38	10.6	162312	2	AC114455	AC114455 Rattus no
42	c	38	10.6	164422	2	AC095588	AC095588 Rattus no
43	c	37.8	10.6	172076	2	AC121895	AC121895 Mus muscu
44		37.8	10.6	194158	2	AC107761	AC107761 Mus muscu
45		37.8	10.6	210824	2	AC126661	AC126661 Rattus no

ALIGNMENTS

RESULT 1
AX039060
LOCUS AX039060 357 bp DNA linear PAT 16-NOV-2000
DEFINITION Sequence 3 from Patent EP1038961.
ACCESSION AX039060
VERSION AX039060.1 GI:11228356
KEYWORDS .
SOURCE Beet necrotic yellow vein virus.
ORGANISM Beet necrotic yellow vein virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
REFERENCE 1 (bases 1 to 357)
AUTHORS Lauber,E., Jonard,G., Guilley,H. and Richards,K.
TITLE Method for inducing viral resistance into a plant
JOURNAL Patent: EP 1038961-A 3 27-SEP-2000;
DE CENTRE NAT (FR)

```
FEATURES             Location/Qualifiers
     source          1. .357
                     /organism="Beet necrotic yellow vein virus"
                     /db_xref="taxon:31721"
     CDS             1. .357
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC16571.1"
                     /db_xref="GI:11228357"
                     /translation="MSREITARPNKNVPIVVGVCVVAFFVLLAFMQQAAATHSGGDYG
                     VPTFSNGGIYRDGTRSADFNSNNHRAYGCGGSGGSVSSRVGQQLIVLAIVSVLIVSLL
                     QRLRSPPEHICNGACG"
BASE COUNT       84 a     60 c     96 g    117 t
ORIGIN

  Query Match           100.0%;  Score 357;  DB 6;  Length 357;
  Best Local Similarity  100.0%;  Pred. No. 3.3e-95;
  Matches  357;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGG 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGG 120

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357


RESULT 2
AX039058
LOCUS       AX039058                357 bp    DNA     linear   PAT 16-NOV-2000
DEFINITION  Sequence 1 from Patent EP1038961.
ACCESSION   AX039058
VERSION     AX039058.1  GI:11228354
KEYWORDS    .
SOURCE      Beet necrotic yellow vein virus.
  ORGANISM  Beet necrotic yellow vein virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
REFERENCE   1  (bases 1 to 357)
  AUTHORS   Lauber,E., Jonard,G., Guilley,H. and Richards,K.
  TITLE     Method for inducing viral resistance into a plant
  JOURNAL   Patent: EP 1038961-A 1 27-SEP-2000;
            DE CENTRE NAT (FR)
FEATURES             Location/Qualifiers
     source          1. .357
                     /organism="Beet necrotic yellow vein virus"
                     /db_xref="taxon:31721"
     CDS             1. .357
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC16570.1"
                     /db_xref="GI:11228355"
                     /translation="MSREITARPNKNVPIVVGVCVVAFFVLLAFMQQKHKTHSGGDYG
                     VPTFSNGGIYRDGTRSADFNSNNHRAYGCGGSGGSVSSRVGQQLIVLAIVSVLIVSLL
                     QRLRSPPEHICNGACG"
BASE COUNT       89 a     58 c     93 g    117 t
ORIGIN

  Query Match           97.3%;  Score 347.4;  DB 6;  Length 357;
  Best Local Similarity  98.3%;  Pred. No. 2.4e-92;
  Matches  351;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGG 120
             |||||||||||||||||||||||||||||||||||  |  |  ||||||||||||
Db      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357


RESULT 3
TOBNYVV2
LOCUS       TOBNYVV2                4612 bp    RNA     linear   VRL 25-MAY-1999
DEFINITION  Beet necrotic yellow vein virus RNA-2.
ACCESSION   X04197
VERSION     X04197.1  GI:62050
KEYWORDS    coat protein; subgenome; unidentified reading frame.
SOURCE      Beet necrotic yellow vein virus.
  ORGANISM  Beet necrotic yellow vein virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
REFERENCE   1  (bases 1 to 4612)
  AUTHORS   Bouzoubaa,S., Ziegler,V., Beck,D., Guilley,H., Richards,K. and
            Jonard,G.
  TITLE     Nucleotide sequence of Beet Necrotic Yellow Vein Virus RNA-2
  JOURNAL   J. Gen. Virol. 67, 1689-1700 (1986)
COMMENT     Partial readthrough of the termination codon at position 709-711
            gives rise to 75K protein
            Data kindly reviewed (02-SEP-1987) by Bouzoubaa S.
FEATURES             Location/Qualifiers
     source          1. .4612
                     /organism="Beet necrotic yellow vein virus"
                     /strain="isolate F13"
                     /db_xref="taxon:31721"
                     /clone="pBF14, pBC2"
     repeat_unit     1. .8
                     /note="inverted repeat A"
     repeat_unit     27. .34
                     /note="inverted repeat A'"
     CDS             145. .2220
                     /note="75K protein (aa 1-691); read-through stop-codon"
                     /codon_start=1
                     /transl_except=(pos:709. .711,aa:OTHER)
                     /protein_id="CAB43511.1"
                     /db_xref="GI:4894189"
                     /translation="MSSEGRYMTWKDMSHNKFMTDRWARVSDVVSVIKQSHAMDLSKA
                     ANLSIIKTALAGLGSGWSDSNPFVSPMTRFPQTLTMYGALVLYVNLSDPEFALIMTKV
                     NTLTDSGLADNASANVRRDVVSGNKAESSGKTAGTNENSAYTLTVSLAGLAQALRLEE
                     LMWTRDKFEDRFKLPWTPVQGRTSPPGQXQLAAARVTAHIRAAKRALLYPGDSPEWVG
                     WKHFYPPPPYDVYDVPPLDIINAKLAADDIGGLVTPTPASSHGLPFEVSEEVEQANRN
                     SLWLTVGLLLAALAIGIGVTAYHRKKLQSRLRELKLLWGSTGGSGGGGGFGTELHTRA
                     TDTVSLGTTLSEHVAPAPSGLRHRPAATDSGPHEVLPFEVWVFDNLAVVYDSIGMSDL
                     FYTVREFVGVFNGEPEGLIELLESPDDDDGVYTNAPRDTAIDAYESQENYDRIDVETV
                     LIERRINLKKLLLEEAELERRERDMTMIADEEQRTLLHRLESSRVEATHVVTKAEADA
```

```
                RAAVAMAALASKEANDYDSKMAFDRSCKEQELRLRELEVNSMPSKTERYVHTGIQGGA
                QLAGAMAVGAMLRRGASSSSQNVSGGANIGSRSQSLTRGRSASQPLSSVGGSTRGVNN
                NNISNTNLVRAGNSAAVSAGRSTNSGNSNFWSKLRVGEGWSKYSVERAATRAQRAIVP
                PAPPSAPAG"
     CDS        145. .711
                /note="coat protein; 21K (aa 1-188)"
                /codon_start=1
                /protein_id="CAA27791.1"
                /db_xref="GI:62051"
                /db_xref="SPTREMBL:Q88620"
                /translation="MSSEGRYMTWKDMSHNKFMTDRWARVSDVVSVIKQSHAMDLSKA
                ANLSIIKTALAGLGSGWSDSNPFVSPMTRFPQTLTMYGALVLYVNLSDPEFALIMTKV
                NTLTDSGLADNASANVRRDVVSGNKAESSGKTAGTNENSAYTLTVSLAGLAQALRLEE
                LMWTRDKFEDRFKLPWTPVQGRTSPPGQ"
     CDS        <712. .2220
                /note="pot. 54K protein (aa 1-503)"
                /codon_start=1
                /protein_id="CAA27792.1"
                /db_xref="GI:1334951"
                /db_xref="SPTREMBL:Q88621"
                /translation="QLAAARVTAHIRAAKRALLYPGDSPEWVGWKHFYPPPPYDVYDV
                PPLDIINAKLAADDIGGLVTPTPASSHGLPFEVSEEVEQANRNSLWLTVGLLLAALAI
                GIGVTAYHRKKLQSRLRELKLLWGSTGGSGGGGGFGTELHTRATDTVSLGTTLSEHVA
                PAPSGLRHRPAATDSGPHEVLPFEVWVFDNLAVVYDSIGMSDLFYTVREFVGVFNGEF
                EGLIELLESPDDDDGVYTNAPRDTAIDAYESQENYDRIDVETVLIERRINLKKLLLEE
                AELERRERDMTMIADEEQRTLLHRLESSRVEATHVVTKAEADARAAVAMAALASKEAN
                DYDSKMAFDRSCKEQELRLRELEVNSMPSKTERYVHTGIQGGAQLAGAMAVGAMLRRG
                ASSSSQNVSGGANIGSRSQSLTRGRSASQPLSSVGGSTRGVNNNNISNTNLVRAGNSA
                AVSAGRSTNSGNSNFWSKLRVGEGWSKYSVERAATRAQRAIVPPAPPSAPAG"
     CDS        2133. .3287
                /note="42K protein (aa 1-384)"
                /codon_start=1
                /protein_id="CAA27793.1"
                /db_xref="GI:62053"
                /db_xref="SPTREMBL:Q89659"
                /translation="MVQVQRRTGGDKGAKGNRASSAPVRSRRMTQDDWSRTHPDDIFS
                VIEKTLVEDGYKWNGVKPGHCDWGKLKESGAIDNFRGTLEGELGKNCDLTCNAAAVKL
                DTLQKVKTSSDWTARVGIVLGAPGVGKSTSIKNLLDKFGAKHKMVLCLPFSQLLEGVF
                AGRLDTFLVDDLFCRSVEYGKYNTMLVDEVTRVHMCEILVLAGHLGVKNVICFGDPAQ
                GLNYKAGSAVNYNFPIIAECYASRRFGKATADLINSSNGGGKPVVGNNEVKDSWTFEE
                LCGKILDMSTVLVATRETQKFLLEDNIESILYSDAHGQTYDVVTIILEDEFDDAAICD
                PNVRAVLLTRARKGGMIKMGPNIAARFKNGDFNSRGVSKSCTGDTFCEDR"
     CDS        3287. .3643
                /note="pot. 13K protein (aa 1-117)"
                /codon_start=1
                /protein_id="CAA27794.1"
                /db_xref="GI:62054"
                /db_xref="SPTREMBL:Q88622"
                /translation="MSREITARPNKNVPIVVGVCVVAFFVLLAFMQQKHKTHSGGDYG
                VPTFSNGGIYRDGTRSADFNSNNHRAYGCGGSGGSVSSRVGQQLIVLAIVSVLIVSLL
                QRLRSPPEHICNGACG"
     CDS        3627. .4025
                /note="pot. 15K protein (aa 1-132)"
                /codon_start=1
                /protein_id="CAA27795.1"
                /db_xref="GI:62055"
                /db_xref="SPTREMBL:Q89500"
                /translation="MVLVVKVDLSNIVLYIVAGCVVVSMLYSPFFSNDVKASSYAGAI
                FKGSGCIMDRNSFAQFGSCDIPKHVAESITKVATKEHDVDIMVKRGEVTVRVVTLTET
                IFIILSRLFGLAVFLFMICLMSIVWFWYHR"
     CDS        4043. .4426
                /note="pot. 14K protein (aa 1-127)"
                /codon_start=1
                /protein_id="CAA27796.1"
                /db_xref="GI:62056"
                /db_xref="SPTREMBL:Q88623"
                /translation="MGMVDSLCVFVGRVITEGSESVEGVERFSIKFSEWKLFTIAVFV
                EYRELGEKECSLKDAGRLHFNVSCVKCCQKLKCKKQNKNHSKHVQNGYLRKVRNFSIL
                GVCGDCCESFTLADEKPHVIVDPEV"
     polyA_site 4612
BASE COUNT     1200 a    760 c   1209 g   1443 t
ORIGIN
```

```
  Query Match             97.3%;  Score 347.4;  DB 14;  Length 4612;
  Best Local Similarity   98.3%;  Pred. No. 2.8e-92;
  Matches  351;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3287 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 3346

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGG 120
          |||||||||||||||||||||||||||||||||||  |  |  ||||||||||||
Db    3347 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 3406

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3407 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 3466

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3467 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 3526

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3527 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 3586

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3587 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 3643
```

```
RESULT 4
BNYVMTP1
LOCUS       BNYVMTP1                1922 bp    RNA     linear   VRL 01-JUN-1994
DEFINITION  Beet necrotic yellow vein mosaic virus (RG1)genomic RNA for
            transport proteins.
ACCESSION   X75575
VERSION     X75575.1  GI:496548
KEYWORDS    13K transport protein; 15K transport protein; 42K transport
            protein; trasnsport protein.
SOURCE      Beet necrotic yellow vein virus.
  ORGANISM  Beet necrotic yellow vein virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
REFERENCE   1  (bases 1 to 1922)
  AUTHORS   Koenig,R.
  JOURNAL   Unpublished
  REMARK    (sites)
REFERENCE   2  (bases 1 to 1922)
  AUTHORS   Solovyev,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-NOV-1993) Solovyev A., Institute for Biochemistry and
            Plant Virology, Messeweg 11/12, Braunschweig, Germany
FEATURES             Location/Qualifiers
     source          1. .1922
                     /organism="Beet necrotic yellow vein virus"
                     /db_xref="taxon:31721"
     CDS             1. .1155
                     /codon_start=1
                     /product="42K transport protein"
                     /protein_id="CAA53253.1"
                     /db_xref="GI:496549"
                     /db_xref="SPTREMBL:Q89659"
                     /translation="MVQVQRRTGGDKGAKGNRASSAPVRSRRMTQDDWSRTHPDDIFS
                     VIEKTLVEDGYKWNGVKPGHCDWGKLKESGAIDNFRGTLEGELGKNCDLTCNAAAVKL
                     DTLQKVKTSSDWTARVGIVLGAPGVGKSTSIKNLLDKFGAKHKMVLCLPFSQLLEGVF
                     AGRLDTFLVDDLFCRSVEYGKYNTMLVDEVTRVHMCEILVLAGHLGVKNVICFGDPAQ
                     GLNYKAGSAVNYNFPIIAECYASRRFGKATADLINSSNGGGKPVVGNNEVKDSWTFEE
                     LCGKILDMSTVLVATRETQKFLLEDNIESILYSDAHGQTYDVVTIILEDEFDDAAICD
                     PNVRAVLLTRARKGGMIKMGPNIAARFKNGDFNSRGVSKSCTGDTFCEDR"
     CDS             1155. .1511
                     /codon_start=1
                     /product="13K transport protein"
                     /protein_id="CAA53254.1"
                     /db_xref="GI:496550"
                     /db_xref="SPTREMBL:Q65682"
```

```
                     /translation="MSREITARPNKNVPIVVGVCVVAFFVLLAFMQQKHKTHSGGDYG
                     VPTFSNGGKYRDGTRSADFNSNNHRAYGCGGSGGSVSSRVGQQLIVLAIVSVLIVSLL
                     QRLRSPPEHICNGACG"
     CDS             1495. .1893
                     /codon_start=1
                     /product="15K transport protein"
                     /protein_id="CAA53255.1"
                     /db_xref="GI:496551"
                     /db_xref="SPTREMBL:Q89500"
                     /translation="MVLVVKVDLSNIVLYIVAGCVVVSMLYSPFFSNDVKASSYAGAI
                     FKGSGCIMDRNSFAQFGSCDIPKHVAESITKVATKEHDVDIMVKRGEVTVRVVTLTET
                     IFIILSRLFGLAVFLFMICLMSIVWFWYHR"
BASE COUNT      505 a    309 c    506 g    602 t
ORIGIN

  Query Match             96.9%;  Score 345.8;  DB 14;  Length 1922;
  Best Local Similarity   98.0%;  Pred. No. 8e-92;
  Matches  350;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1155 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 1214

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGG 120
          ||||||||||||||||||||||||||||||||||||  |   |  ||||||||||||||
Db    1215 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 1274

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    1275 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTAAATATAGAGACGGTACAAGATCA 1334

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1335 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 1394

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1395 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 1454

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1455 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 1511


RESULT 5
E59890
LOCUS       E59890                   739 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Rhizomania-resisting plant.
ACCESSION   E59890
VERSION     E59890.1  GI:18622726
KEYWORDS    JP 2000312540-A/2.
SOURCE      Beet necrotic yellow vein virus.
  ORGANISM  Beet necrotic yellow vein virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
REFERENCE   1  (bases 1 to 739)
  AUTHORS   Nomura,S., Kamitani,G., Saito,M., Kiguchi,T., Kusu,S. and Soma,C.
  TITLE     Rhizomania-resisting plant
  JOURNAL   Patent: JP 2000312540-A 2 14-NOV-2000;
            GENICHI KAMITANI,SHADANHOJIN HOKKAIDO TENSAI KYOKAI
COMMENT     OS   Beet Necrotic Yellow Vein Virus
            PN   JP 2000312540-A/2
            PD   14-NOV-2000
            PF   28-APR-1999 JP 1999122628
            PR
            PI   SHINJI NOMURA,GENICHI KAMITANI,MINAKO SAITO,TADAHIKO KIGUCHI,
            PI   SHUNZO KUSUME,
            PI   CHIHIRO SOMA
            PC   A01H5/00,C12N5/10,C12N15/09,C12N5/00,C12N15/00 CC
            FH   Key             Location/Qualifiers
            FT   source          1. .739
            FT                   /organism='Beet Necrotic Yellow Vein Virus'.
FEATURES             Location/Qualifiers
     source          1. .739
                     /organism="Beet necrotic yellow vein virus"
                     /db_xref="taxon:31721"
BASE COUNT      184 a    114 c    189 g    252 t
ORIGIN

  Query Match             91.9%;  Score 328.2;  DB 6;  Length 739;
  Best Local Similarity   95.0%;  Pred. No. 1.3e-86;
  Matches  339;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGG 120
          ||||||||||||||||||||||||||||||||||||  |   |  ||||||||||||||
Db      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
          || ||||||||||||||||||||||||||||||||| ||| |||||||||||||| |||
Db     121 GGCGATTACGGAGTCCCAACATTTTCTAACGGTGGTAAATACAGAGACGGTACAAGGTCA 180

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240
          |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db     181 GCTGATTTTAATAGTAATAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
          |||||||||| |||||||||||| |||||||||||||||| |||||||||||||| |||
Db     241 AGTAGTCGAGTCGGGCAGCAACTTGTTGTGTTAGCTATTGTGTCTGTGTTAATAGTATCA 300

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
          || ||||||||||||| |||||||||||||| |||||||||||||||||||||||||
Db     301 CTGTTACAACGATTAAGATCTCCACCAGAACATATTTGTAATGGTGCTTGTGGTTAA 357


RESULT 6
AF197556
LOCUS       AF197556                 4544 bp    RNA     linear   VRL 04-DEC-2000
DEFINITION  Beet necrotic yellow vein virus isolate Kas3 RNA 2 segment, partial
            sequence.
ACCESSION   AF197556
VERSION     AF197556.1  GI:11528045
KEYWORDS    .
SOURCE      Beet necrotic yellow vein virus.
  ORGANISM  Beet necrotic yellow vein virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
REFERENCE   1  (bases 1 to 4544)
  AUTHORS   Koenig,R. and Lennefors,B.L.
  TITLE     Molecular analyses of European A, B and P type sources of Beet
            necrotic yellow vein virus and detection of the rare P type in
            Kazakhstan
  JOURNAL   Arch. Virol. 145 (8), 1561-1570 (2000)
  MEDLINE   20456802
   PUBMED   11003469
REFERENCE   2  (bases 1 to 4544)
  AUTHORS   Koenig,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-1999) Institut fuer Pflanzenvirologie,
            Mikrobiologie und Biologische Sicherheit, Biologische Bundesanstalt
            fuer Land und Forstwirtschaft, Messeweg 11, D 38104 Braunschweig D
            38104, Germany
FEATURES             Location/Qualifiers
     source          1. .4544
                     /organism="Beet necrotic yellow vein virus"
                     /isolate="Kas3"
                     /db_xref="taxon:31721"
                     /country="Kazakhstan"
                     /note="closely related to P type
                     almost complete sequence of RNA 2; lacks only a few
                     nucleotides at the 5'-  and 3'-ends
                     type: P; segment: RNA 2"
     CDS             join(107. .670,674. .2179)
                     /note="contains read-through stop codon"
                     /codon_start=1
```

```
                     /product="coat protein"
                     /protein_id="AAG37092.1"
                     /db_xref="GI:11528046"
                     /translation="MSSEGRYMTWKDMSHNRFMTDRWARVSDVVSVIKQSHAMDLSKA
                     ANLSIIKTALAGLGSGWTDNNPFVSPMTRFPQTLTMYGALVLYVNLSDPEFALIMTKI
                     STLTDSGLADNASANVRRDVVSGNKAESSGKTAGTNENSAYTLTVSLAGLAQALRLEE
                     LMWTRDKFEDRLKLPWTPVQGRTSPPGQQLAAARVTAHIRAAKRALLYPGDSPEWVGW
                     KHFYPPPPYDVYDVPPLDIINAKLAADDIGGLVTPTPASSHGLPFEVSEEVEQANRNS
                     LWLTVGLLLAALAVGIGVAAYHRKKLQSRLRELKLLWGSTGGSGGGGGFDTELYMRAT
                     DTVSLGTTLSEHAASAPSGLRHRPAATDSGPHEALPFEVWAFDNLSVVYDSIGMSDLF
                     YTVREFVGVFNGEFEGLIELLESPDDDDGVYTNAPRDTAIDAYESQENYDRIDIETVL
                     LERRINLKKLLLEEAELERRERDMTMIADEEQRTLLHRLESSRVEATHAVAKAEADAR
                     AAVAMTALASKEANDYDSKMAFDRSCKEQELRLRELEVNSMPSKTERYVHTGIQGGAQ
                     LAGAMAVGAMLRRGASSSSQTVSSGANIGSRSQSVTRGRSASQPLSSVGGFTRGVNNN
                     ISNTNLVRAGNSAEVSAGRSTNSGNSNFWSKLRVGEGWSKYSVERAATRAQRAIMLPA
                     PPSAPAG"
     misc_feature    671. .673
                     /note="read-through stop codon"
     CDS             2092. .3246
                     /note="first triple gene block; necessary for virus
                     movement"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAG37093.1"
                     /db_xref="GI:11528047"
                     /translation="MVQVQRRTGGDKGAKGNYASSAPVRSRRMTQDDWSRTHPDDIFS
                     VIEKTLVEDGYKWNGVKPGHCDWGKLKESGAIDNFRGTLEGELGKNCDLTCNAAAVKL
                     DTLQKVKMSSDWTARVGIVLGAPGVGKSTSIKNLLDKFGAKHKMVLCLPFSQLLEGVF
                     AGRLDTFLVDDLFCRSVEYGKYNTMLVDEVTRVHMCEILVLAGHLGVKNVICFGDPAQ
                     GLNYKAGSAVNYNFPIIAECYASRRFGKATADLINSSNGGGKPVVGNNEVKDSWTFEE
                     LCGKILDMSTVLVATRETQKFLLEDNIESILYSDAHGQTYDVVTIILEDEFDDAAICD
                     PNVRAVLLTRARKGGMIKVGPNIAARFKNGDFNSRGVSKSCTGDTFCEDR"
     CDS             3246. .3602
                     /note="second triple gene block; necessary for virus
                     movement"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAG37094.1"
                     /db_xref="GI:11528048"
                     /translation="MSREITARPNKNVPIVVGVCVVAFFVLLAFMQQKHKTHSGGDYG
                     VPTFSNGGKYRDGTRSADFNSNNHRAYGCGGSGGSVSSRVGQQLVVLAIVSVLIVSLL
                     QRLRSPPEHICNGACG"
     CDS             3586. .3984
                     /note="third triple gene block; necessary for virus
                     movement"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAG37095.1"
                     /db_xref="GI:11528049"
                     /translation="MVLVVKVDLSNIVLYIVAGCVVVSMLYSPFFSNDVKASSYAGAV
                     FRGSGCIMDRNSFAQFGSCDIPKHVAESITKVATKEHDADVMAKRGEVTVRVVTLTET
                     IFIILSRLFGLAVFLFMICLMSIVWFWCHR"
     CDS             4002. .4385
                     /codon_start=1
                     /product="P14 protein"
                     /protein_id="AAG37096.1"
                     /db_xref="GI:11528050"
                     /translation="MGMVDSLCVFVGRVITEGSESVEGVERFSIKFSEWKLLTTAVYV
                     EYRQLGEKECSLKDVGRLHFNMSCVKCCRKLKCEKQNKNHSKHVQNGYLRKVRNFSIL
                     GVCGDCCESFTLADEKHHVIVDPEV"
BASE COUNT     1174 a    756 c   1207 g   1407 t
ORIGIN

  Query Match            91.9%;  Score 328.2;  DB 14;  Length 4544;
  Best Local Similarity  95.0%;  Pred. No. 1.4e-86;
  Matches  339;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3246 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 3305

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGG 120
          |||||||||||||||||||||||||||||||||||||  |  |  ||||||||||||
Db    3306 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 3365

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
          || ||||||||||||||||||||||||||||||||||| |||||||||||||||| |||
Db    3366 GGCGATTACGGAGTCCCAACATTTTCTAACGGTGGTAAATATAGAGACGGTACAAGGTCA 3425

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 240
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    3426 GCTGATTTTAATAGTAATAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 3485

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
          |||||||||| |||||||||||| |||||||||||||||| |||||||| ||||| |||
Db    3486 AGTAGTCGAGTCGGGCAGCAACTTGTTGTGTTAGCTATTGTGTCTGTGTTGATAGTATCA 3545

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
          || |||||||||||||| |||||||||||||| |||||||||||||||||||||||
Db    3546 CTGTTACAACGATTAAGATCTCCACCAGAACATATTTGTAATGGTGCTTGTGGTTAA 3602


RESULT 7
E59889
LOCUS       E59889                  4609 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Rhizomania-resisting plant.
ACCESSION   E59889
VERSION     E59889.1  GI:18622725
KEYWORDS    JP 2000312540-A/1.
SOURCE      Beet necrotic yellow vein virus.
  ORGANISM  Beet necrotic yellow vein virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
REFERENCE   1  (bases 1 to 4609)
  AUTHORS   Nomura,S., Kamitani,G., Saito,M., Kiguchi,T., Kusu,S. and Soma,C.
  TITLE     Rhizomania-resisting plant
  JOURNAL   Patent: JP 2000312540-A 1 14-NOV-2000;
            GENICHI KAMITANI,SHADANHOJIN HOKKAIDO TENSAI KYOKAI
COMMENT     OS   Beet necrotic yellow vein virus
            PN   JP 2000312540-A/1
            PD   14-NOV-2000
            PF   28-APR-1999 JP 1999122628
            PR
            PI   SHINJI NOMURA,GENICHI KAMITANI,MINAKO SAITO,TADAHIKO KIGUCHI,
            PI   SHUNZO KUSUME,
            PI   CHIHIRO SOMA
            PC   A01H5/00,C12N5/10,C12N15/09,C12N5/00,C12N15/00 CC
            FH   Key             Location/Qualifiers
            FT   source          1. .4609
            FT                   /organism='Beet necrotic yellow vein virus'.
FEATURES             Location/Qualifiers
     source          1. .4609
                     /organism="Beet necrotic yellow vein virus"
                     /db_xref="taxon:31721"
BASE COUNT     1187 a    765 c   1221 g   1436 t
ORIGIN

  Query Match            91.9%;  Score 328.2;  DB 6;  Length 4609;
  Best Local Similarity  95.0%;  Pred. No. 1.4e-86;
  Matches  339;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3284 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 3343

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGG 120
          |||||||||||||||||||||||||||||||||||||  |  |  ||||||||||||
Db    3344 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 3403

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
          || ||||||||||||||||||||||||||||||||| ||| |||||||||||||| |||
Db    3404 GGCGATTACGGAGTCCCAACATTTTCTAACGGTGGTAAATACAGAGACGGTACAAGGTCA 3463

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 240
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    3464 GCTGATTTTAATAGTAATAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 3523
```

```
Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
           |||||||||| |||||||||| |||||||||||||| |||||||||||| |||
Db    3524 AGTAGTCGAGTCGGGCAGCAACTTGTTGTGTTAGCTATTGTGTCTGTGTTAATAGTATCA 3583

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
           || |||||||||||| |||||||||||| |||||||||||||||||||||||||
Db    3584 CTGTTACAACGATTAAGATCTCCACCAGAACATATTTGTAATGGTGCTTGTGGTTAA 3640


RESULT 8
BNYKPRO2
LOCUS       BNYKPRO2                4609 bp    RNA     linear   VRL 06-FEB-1999
DEFINITION  Beet necrotic yellow vein virus genomic RNA for 42k protein, 13k
            protein, 15k protein, 14k protein, 75k protein.
ACCESSION   D84411
VERSION     D84411.1  GI:1304018
KEYWORDS    coat protein; pot. 54k protein; 75k protein; 14k protein; 15k
            protein; 13k protein; 42k protein.
SOURCE      Beet necrotic yellow vein virus (isolate:S, lab_host:Tetragonia
            expansa) cDNA to genomic RNA, clone:145-1, H20, S4, S7, pMPCRD5,
            pMPCRSX19, pMRP70, pMSD25.
  ORGANISM  Beet necrotic yellow vein virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
REFERENCE   1  (sites)
  AUTHORS   Saito,M., Kiguchi,T., Kusume,T. and Tamada,T.
  TITLE     Complete nucleotide sequence of the Japanese isolate S of beet
            necrotic yellow vein virus RNA and comparison with European
            isolates
  JOURNAL   Arch. Virol. 141 (11), 2163-2175 (1996)
  MEDLINE   97128991
REFERENCE   2  (bases 1 to 4609)
  AUTHORS   Saito,M., Kiguchi,T., Kusume,T. and Tamada,T.
  TITLE     Complete nucleotide sequence of a Japanese isolate S of beet
            necrotic yellow vein virus and comparison with European isolates
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 4609)
  AUTHORS   Kiguchi,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-1996) Tadahiko Kiguchi, Hokkaido Central
            Agricultural Experiment Station, Plant Biotechnology; Higashi 6
            Kita 15, Naganuma, Hokkaido 069-13, Japan (Tel:01238-9-2001,
            Fax:01238-9-2060)
FEATURES             Location/Qualifiers
     source          1. .4609
                     /organism="Beet necrotic yellow vein virus"
                     /isolate="S"
                     /db_xref="taxon:31721"
                     /clone="145-1, H20, S4, S7, pMPCRD5, pMPCRSX19, pMRP70,
                     pMSD25"
                     /lab_host="Tetragonia expansa"
     5'UTR           1. .144
     CDS             join(145. .708,712. .2217)
                     /note="readthrough"
                     /codon_start=1
                     /product="75k protein"
                     /protein_id="BAA18968.1"
                     /db_xref="GI:1902910"
                     /translation="MSSEGRYMTWKDMSHNKFMTDRWARVSDVVSVIKQSHAMDLSKA
                     ANLSIIKTALAGLGSGWTDNNPFVSPMTRFPQTLTMYGALVLYVNLSDPEFALIMTKV
                     STLTDSGLADNASANVRRDVVSGNKAESSGKTAGTNENSAYTLTVSLAGLAQALRLEE
                     LMWTRDKFEDRLKLPWTPVQGRTSPPGQQLAAARVTAHIRAAKRALLYPGDSPEWVGW
                     KHFYPPPPYDVYDVPPLDIINAKLAADDIGGLVTPTPASSHGLPFEVSEEVEQANRNS
                     LWLTVGLLLAALAVGIGVAAYHRKKLQSRLRELKLLWGSTGGSGGGGGFDTELYMRAT
                     DTVSLGTTLSEHAASAPSGLRHRPAATDSGPHEALPFEVWVFDNLAVVYDSIGMSDLF
                     YTVREFVGVFNGEFEGLIELLESPDDDDGVYTNAPRDTAIDAYESQENYDRIDIETVL
                     IERRINLKKLLLEEAELERRERDMTMIADEEQRTLLHRLESSRVEATHAVAKAEADAR
                     AAVAMAALASKEANDYDSKMAFDRSCKEQELRLRELEVNSMPSKTERYVHTGIQGGAQ
                     LAGAMAVGAMLRRGAGSSSQTVSSGANIGSRSQSLTRGRSASQPLSSVGGSTRGVNNN
                     ISNTNLVRAGNSAEVSAGRSTNSGNSNFWSKLRVGEGWSKYSVERAATRAQRAIVLPA
                     PPSAPAG"
     mat_peptide     145. .708
                     /product="coat protein"
     mat_peptide     712. .2214
                     /product="pot. 54k protein"
     CDS             2130. .3284
                     /codon_start=1
                     /product="42k protein"
                     /protein_id="BAA12342.1"
                     /db_xref="GI:1304022"
                     /translation="MVQVQRRTGGDKGAKGNRASSAPVRSRRMTQDDWSRTHPDDIFS
                     VIEKTLVEDGYKWNGVKPGHCDWGKLKESGAIDNFRGTLEGELGKNCDLTCNAAAVKL
                     DTLQKVKMSSDWTARVGIVLGAPGVGKSTSIKNLLDKFGAKHKMVLCLPFSQLLEGVF
                     AGRLDTFLVDDLFCRSVGYGKYNTMLVDEVTRVHMCEILVLAGHLGVKNVICFGDPAQ
                     GLNYKAGSAVNYNFPIIAECYASRRFGKATADLINSSNGGGKPVVGNNEVKDSWTFEE
                     LCGKILDMSTVLVATRETQKFLLEDNIESILYSDAHGQTYDVVTIILEDEFDDAAICD
                     PNVRAVLLTRARKGGMIKMGPNIAARFKNGDFNSRGVSKSCTGDTFCEDR"
     CDS             3284. .3640
                     /codon_start=1
                     /product="13k protein"
                     /protein_id="BAA12343.1"
                     /db_xref="GI:1304023"
                     /translation="MSREITARPNKNVPIVVGVCVVAFFVLLAFMQQKHKTHSGGDYG
                     VPTFSNGGKYRDGTRSADFNSNNHRAYGCGGSGGSVSSRVGQQLVVLAIVSVLIVSLL
                     QRLRSPPEHICNGACG"
     CDS             3624. .4022
                     /codon_start=1
                     /product="15k protein"
                     /protein_id="BAA12344.1"
                     /db_xref="GI:1304024"
                     /translation="MVLVVKVDLSNIVLYIVAGCVVVSMLYSPFFSNDVKASSYAGAV
                     FKGSGCIMDRNSFAQFGSCDIPKHVAESITKVATKEHDADIMVKRGEVTVRVVTLTET
                     LFIILSRLFGLAVFLFMICLMSIVWFWCHR"
     CDS             4034. .4423
                     /codon_start=1
                     /product="14k protein"
                     /protein_id="BAA12345.1"
                     /db_xref="GI:1304025"
                     /translation="MSMGMVDSLCVFVGRVITEGSESVEGVERFSIKFSDWKLFTTAV
                     YVEYRQLGEKECSLKDVGRLHFNMSCVKCCQKLKCKKQNKNHSKHVQNGYLRKVRNFS
                     ILGVCGDCCESFTLADEKHHVIVDPEV"
BASE COUNT     1187 a    765 c   1221 g   1436 t
ORIGIN

  Query Match             91.9%;  Score 328.2;  DB 14;  Length 4609;
  Best Local Similarity   95.0%;  Pred. No. 1.4e-86;
  Matches  339;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3284 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 3343

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGG 120
           |||||||||||||||||||||||||||||||||||  | |  ||||||||||||||
Db    3344 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 3403

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
           || ||||||||||||||||||||||||||||||| ||| ||||||||||||| |||
Db    3404 GGCGATTACGGAGTCCCAACATTTTCTAACGGTGGTAAATACAGAGACGGTACAAGGTCA 3463

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 240
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    3464 GCTGATTTTAATAGTAATAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 3523

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
           |||||||||| |||||||||||| |||||||||||||| |||||||||||| |||
Db    3524 AGTAGTCGAGTCGGGCAGCAACTTGTTGTGTTAGCTATTGTGTCTGTGTTAATAGTATCA 3583

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
           || |||||||||||||| |||||||||||||| |||||||||||||||||||||||||
Db    3584 CTGTTACAACGATTAAGATCTCCACCAGAACATATTTGTAATGGTGCTTGTGGTTAA 3640


RESULT 9
AF197547
LOCUS       AF197547                4544 bp    RNA     linear   VRL 04-DEC-2000
DEFINITION  Beet necrotic yellow vein virus isolate F75 RNA 2 segment, partial
```

```
            sequence.
ACCESSION   AF197547
VERSION     AF197547.1  GI:11528023
KEYWORDS    .
SOURCE      Beet necrotic yellow vein virus.
  ORGANISM  Beet necrotic yellow vein virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
REFERENCE   1  (bases 1 to 4544)
  AUTHORS   Koenig,R. and Lennefors,B.L.
  TITLE     Molecular analyses of European A, B and P type sources of Beet
            necrotic yellow vein virus and detection of the rare P type in
            Kazakhstan
  JOURNAL   Arch. Virol. 145 (8), 1561-1570 (2000)
  MEDLINE   20456802
   PUBMED   11003469
REFERENCE   2  (bases 1 to 4544)
  AUTHORS   Koenig,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-1999) Institut fuer Pflanzenvirologie,
            Mikrobiologie und Biologische Sicherheit, Biologische Bundesanstalt
            fuer Land und Forstwirtschaft, Messeweg 11, D 38104 Braunschweig D
            38104, Germany
FEATURES             Location/Qualifiers
     source          1. .4544
                     /organism="Beet necrotic yellow vein virus"
                     /isolate="F75"
                     /db_xref="taxon:31721"
                     /country="France"
                     /note="almost complete sequence of RNA 2; lacks only a few
                     nucleotides at the 5'-  and 3'-ends
                     type: P; segment: RNA 2"
     CDS             join(107. .670,674. .2179)
                     /note="contains read-through stop codon"
                     /codon_start=1
                     /product="coat protein"
                     /protein_id="AAG37079.1"
                     /db_xref="GI:11528024"
                     /translation="MSSEGRYMTWKDMSHNRFMTDRWARVSDVVSVIKQSHAMDSSKA
                     ANLSIIKTALAGLGSGWTDNNPFVSPMTRFPQTLTMYGALVLYVNLSDPEFALIMTKI
                     STLTDSGLADNASANVRRDVVSGNKAESSGKTAGTNENSAYTLTVSLAGLAQALRLEE
                     LMWTRDKFEDRLKLPWTPVQGRTSPPGQQLAAARVTAHIRAAKRALLYPGDSPEWVGW
                     KHFYPPPPYDVYDVPPLDIINAKLAADDIGGLVTPTPASSHGLPFEVSEEVEQANRNS
                     LWLTVGLLLAALAVGIGVAAYHRKKLQSRLRELKLLWGSTGGSGGGGGLDTELYMRAT
                     DTVSLGTTPSEHAALAPSGLRHRPAATDSGPHEALPFEVWVFDNLAVVYDSIGMSDLF
                     YTVREFVGVFNGEFEGLIELFESPDDDDGVYTNAPRDTAIDAYESQENYDRIDIETVL
                     LERRINLKKLLLEEAELERRERDMTMIADEEQRTLLHRLESSRVEAAHAVAKAEADAR
                     AAVAMAALASKEANDYDSKMAFDRSCKEQELRLRELEVNSMPSKTERYVHTGIQGGAQ
                     LAGAMAVGAMLRRGASSSSQTVSSGANIGSRSQSLTRGRSASQPLSSVGGSTRGVNNN
                     ISNTNLVRAGNSAEVSAGRSTNSGNSNFWSKLRVGEGWSKYSVERAATRAQRAITLPA
                     PPSAPAG"
     misc_feature    671. .673
                     /note="read-through stop codon"
     CDS             2092. .3246
                     /note="first triple gene block; necessary for virus
                     movement"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAG37080.1"
                     /db_xref="GI:11528025"
                     /translation="MVQVQRRTGGDKGAKGNYASSAPVRSRRMTQDDWSRTHPDDIFS
                     VIEKTLVEDGYKWNGVKPGHCDWGKLKESGAIDNFRGTLEGELGKNCDLPCNAAAVKL
                     DTLQKVKMSSDWTARVGIVLGAPGVGKSTSIKNLLDKFGAKHKMVLCLPFSQLLEGVF
                     AGRLDTFLVDDLFCRSVEYGKYNTMLVDEVTRVHMCEILVLAGHLGVKNVICFGDPAQ
                     GLNYKAGSAVNYNFPIIAECYASRRFGKATADLINSSNGGGKPVVGNNEVKDSWTFEE
                     LCGKILDMSTVLVATRETQKFLLEDNIESILYSDAHGQTYDVVTIILEDEFDDAAICD
                     PNVRAVLLTRARKGGMIKMGPNIAARFKNGDFNSRGVSKSCTGDTFCEDR"
     CDS             3246. .3602
                     /note="second triple gene block; necessary for virus
                     movement"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAG37081.1"
                     /db_xref="GI:11528026"
                     /translation="MSREITARPNKNVPIVVGVCVVAFFVLLAFMQQKHKTHSGGDYG
                     VPTFSNGGKYRDRTRSADFNSNNHRAYGCGGSGGSVSSRVGQQLVVLAIVSVLIVSLL
                     QRLRSPPEHICNGACG"
     CDS             3586. .3984
                     /note="third triple gene block; necessary for virus
                     movement"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAG37082.1"
                     /db_xref="GI:11528027"
                     /translation="MVLVVKVDLSNIVLYIVAGCVVVSMLYSPFFSNDVKASSYAGAV
                     FRGSGCIMDRNSFAQFGSCDIPKHVAESITKVATKEHDADVMVKRGEVTVRVVTPTET
                     IFIILSRLFGLAVFLFMICLMSIVWFWCHR"
     CDS             4002. .4385
                     /codon_start=1
                     /product="P14 protein"
                     /protein_id="AAG37083.1"
                     /db_xref="GI:11528028"
                     /translation="MGMVDSLCVFVGRVITEGSESVEGVERFSIKFSEWKLFTTAVYV
                     EYRQLGEKECSLKDVGRLHFNMSCVKCCRKLKCKKQNKNHSKHVQNGYLRKVRNFSIL
                     GVCGDCCESFTLADEKHHVIVDPEV"
BASE COUNT     1174 a    766 c   1205 g   1399 t
ORIGIN

  Query Match             90.6%;  Score 323.4;  DB 14;  Length 4544;
  Best Local Similarity   94.1%;  Pred. No. 3.9e-85;
  Matches  336;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    3246 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTCTGT 3305

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGG 120
             |||||||||||||||||||||||||||||||||||  |  |  ||||||||||||
Db    3306 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 3365

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
             || |||||||||||||||||||||||||||||||||| |||||||||| | ||||| |||
Db    3366 GGCGATTACGGAGTCCCAACATTTTCTAACGGTGGTAAATATAGAGACCGCACAAGGTCA 3425

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 240
             |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db    3426 GCTGATTTTAATAGTAATAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 3485

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
             |||||||||||| |||||||||||| |||||||||||||||| |||||||| |||| |||
Db    3486 AGTAGTCGAGTCGGGCAGCAACTTGTTGTGTTAGCTATTGTGTCTGTGTTGATAGTATCA 3545

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
             || ||||||||||||| |||||||||||||| |||||||||||||||||||||||
Db    3546 CTGTTACAACGATTAAGATCTCCACCAGAACATATTTGTAATGGTGCTTGTGGTTAA 3602


RESULT 10
BNYVMTP2
LOCUS       BNYVMTP2                1922 bp    RNA     linear   VRL 01-AUG-1994
DEFINITION  Beet necrotic yellow vein mosaic virus (Yu2) genomic RNA for
            transport proteins.
ACCESSION   X75574
VERSION     X75574.1  GI:496552
KEYWORDS    13K transport protein; 15K transport protein; 42K transport
            protein; transport protein.
SOURCE      Beet necrotic yellow vein virus.
  ORGANISM  Beet necrotic yellow vein virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
REFERENCE   1  (bases 1 to 1922)
  AUTHORS   Kruse,M., Koenig,R., Hoffmann,A., Kaufmann,A., Commandeur,U.,
            Solovyev,A.G., Savenkov,I. and Burgermeister,W.
  TITLE     Restriction fragment length polymorphism analysis of reverse
            transcription-PCR products reveals the existence of two major
            strain groups of beet necrotic yellow vein virus
  JOURNAL   J. Gen. Virol. 75 (Pt 8), 1835-1842 (1994)
  MEDLINE   94321971
```

```
   PUBMED   7913953
REFERENCE   2  (bases 1 to 1922)
  AUTHORS   Solovyev,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-NOV-1993) Solovyev A., Institute for Biochemistry and
            Plant Virology, Messeweg 11/12, Braunschweig, Germany
FEATURES             Location/Qualifiers
     source          1. .1922
                     /organism="Beet necrotic yellow vein virus"
                     /db_xref="taxon:31721"
     CDS             1. .1155
                     /codon_start=1
                     /product="42K transport protein"
                     /protein_id="CAA53250.1"
                     /db_xref="GI:496553"
                     /db_xref="SPTREMBL:Q65683"
                     /translation="MVQVQRRTGGDKGAKGNRASSAPVRSRRMTQDDWSRTHPDDIFS
                     VIEKTLVEDGYKWNGVKPGHCDWGKLKESGAVDNFRGTLEGELGKNCDLTCNAAAVKL
                     DTLQKVKMSSDWTARVGIVLGAPGVGKSTSIKNLLDKFGAKHKMVLCLPFSQLLEGVF
                     AGRLDTFLVDDLFCRSVEYGKYNTMLVDEVTRVHMCEILVLAGHLGVKNVICFGDPAQ
                     GLNYKAGSAVNYNFPIIAECYASRRFGKATADLINSSNGGGKPVVGNNEVKDSWTFEE
                     LCGKILDMSTVLVATRETQKFLLEDNIESILYSDAHGQTYDVVTIILEDEFDDAAICD
                     PNVRAVFLTRARKGGMIKMGPNIAARFKNGDFNSRGVSKSCTGDTFCEDR"
     CDS             1155. .1511
                     /codon_start=1
                     /product="13K transport protein"
                     /protein_id="CAA53251.1"
                     /db_xref="GI:496554"
                     /db_xref="SPTREMBL:Q65684"
                     /translation="MSREITARPNKNVPIVVGVCVVAFFVLLAFMQQKHKTHAGGDYG
                     VPTFSNGGKYRDGTRSADFNSNNHRAYGCGGSGGSVSSRVGQQLVVLAIVSVVIVLLL
                     QRLRSPPEHICNGACG"
     CDS             1495. .1893
                     /codon_start=1
                     /product="42K transport protein"
                     /protein_id="CAA53252.1"
                     /db_xref="GI:496555"
                     /db_xref="SPTREMBL:Q65685"
                     /translation="MVLVVKVDLSNIVLYIVAGCVVVSMLYSPFFSNDVKASSYAGAV
                     FKGSGCIMDRNSFAQFGSCDIPKHVAESITKVATKEHDADIMVKRGDVTVRVVTLTET
                     IFIILSRLFGLAVFLFMICLMSIVWFWCHR"
BASE COUNT      505 a    309 c    509 g    599 t
ORIGIN

  Query Match             90.1%;  Score 321.8;  DB 14;  Length 1922;
  Best Local Similarity   93.8%;  Pred. No. 1.1e-84;
  Matches  335;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    1155 ATGTCTAGGGAAATAACCGCCCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 1214

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGG 120
          |||||||||||||||||||||||||||||||||||||  |  |  ||||||| |||||
Db    1215 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATGCTGGG 1274

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
          || ||||||||||| |||||||||||||||||||| |||||||||||||||||||| |||
Db    1275 GGCGATTACGGAGTTCCAACATTTTCTAACGGTGGTAAATATAGAGACGGTACAAGGTCA 1334

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    1335 GCTGATTTTAATAGTAATAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 1394

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
          |||||||||| |||||||||||| |||||||||||||||| |||||| ||||||| | |
Db    1395 AGTAGTCGAGTCGGGCAGCAACTTGTTGTGTTAGCTATTGTGTCTGTGGTAATAGTATTA 1454

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
          || |||||||||||||| |||||||||||||| ||||||||||||||||||||||||
Db    1455 CTGTTACAACGATTAAGATCTCCACCAGAACATATTTGTAATGGTGCTTGTGGTTAA 1511
```

```
RESULT 11
E59891
LOCUS       E59891                   739 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Rhizomania-resisting plant.
ACCESSION   E59891
VERSION     E59891.1  GI:18622727
KEYWORDS    JP 2000312540-A/3.
SOURCE      Beet necrotic yellow vein virus.
  ORGANISM  Beet necrotic yellow vein virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
REFERENCE   1  (bases 1 to 739)
  AUTHORS   Nomura,S., Kamitani,G., Saito,M., Kiguchi,T., Kusu,S. and Soma,C.
  TITLE     Rhizomania-resisting plant
  JOURNAL   Patent: JP 2000312540-A 3 14-NOV-2000;
            GENICHI KAMITANI,SHADANHOJIN HOKKAIDO TENSAI KYOKAI
COMMENT     OS   Beet Necrotic Yellow Vein Virus
            PN   JP 2000312540-A/3
            PD   14-NOV-2000
            PF   28-APR-1999 JP 1999122628
            PR
            PI   SHINJI NOMURA,GENICHI KAMITANI,MINAKO SAITO,TADAHIKO KIGUCHI,
            PI   SHUNZO KUSUME,
            PI   CHIHIRO SOMA
            PC   A01H5/00,C12N5/10,C12N15/09,C12N5/00,C12N15/00 CC
            FH   Key             Location/Qualifiers
            FT   mutation        (151). .(174).
FEATURES             Location/Qualifiers
     source          1. .739
                     /organism="Beet necrotic yellow vein virus"
                     /db_xref="taxon:31721"
BASE COUNT      182 a    117 c    187 g    253 t
ORIGIN

  Query Match             88.3%;  Score 315.4;  DB 6;  Length 739;
  Best Local Similarity   92.7%;  Pred. No. 8.1e-83;
  Matches  331;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGG 120
          |||||||||||||||||||||||||||||||||||||  |  |  ||||||||||||||
Db      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
          || ||||||||||||||||||||||||||| || || ||  |||| |  |||| |||
Db     121 GGCGATTACGGAGTCCCAACATTTTCTAACGCTGCTAAATTCAGAGCTGCAGCAAGGTCA 180

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db     181 GCTGATTTTAATAGTAATAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
          |||||||||| |||||||||||| |||||||||||||||| |||||||||||||||| |||
Db     241 AGTAGTCGAGTCGGGCAGCAACTTGTTGTGTTAGCTATTGTGTCTGTGTTAATAGTATCA 300

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
          || |||||||||||||| |||||||||||||| ||||||||||||||||||||||||
Db     301 CTGTTACAACGATTAAGATCTCCACCAGAACATATTTGTAATGGTGCTTGTGGTTAA 357
```

```
RESULT 12
AF061869
LOCUS       AF061869                 4616 bp    RNA     linear   VRL 26-MAR-2002
DEFINITION  Beet soil-borne mosaic virus RNA2, complete genome.
ACCESSION   AF061869
VERSION     AF061869.1  GI:3136264
KEYWORDS    .
SOURCE      beet soil-borne mosaic virus.
  ORGANISM  beet soil-borne mosaic virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.
REFERENCE   1  (bases 1 to 4616)
```

```
  AUTHORS   Lee,L., Telford,E.B., Batten,J.S., Scholthof,K.B. and Rush,C.M.
  TITLE     Complete nucleotide sequence and genome organization of Beet
            soilborne mosaic virus, a proposed member of the genus Benyvirus
  JOURNAL   Arch. Virol. 146 (12), 2443-2453 (2001)
  MEDLINE   21670014
   PUBMED   11811691
REFERENCE   2  (bases 1 to 4616)
  AUTHORS   Batten,J.S., Rush,C.M. and Scholthof,K.-B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-APR-1998) Plant Pathology and Microbiology, Texas A&M
            University, 120 Peterson Building, College Station, TX 77843-2132,
            USA
FEATURES             Location/Qualifiers
     source          1. .4616
                     /organism="beet soil-borne mosaic virus"
                     /strain="EA"
                     /db_xref="taxon:76343"
                     /note="RNA2; Furovirus; submitter suggested
                     classification: genus Benevirus, family Tubiviridae"
     misc_feature    1
                     /note="5'cap"
     CDS             146. .2206
                     /note="translational readthrough of the 21 kDa coat
                     protein gene; sequence contains an amber codon at amino
                     acid (Tyr) position 192, which is a putative readthrough
                     codon (UAG)"
                     /codon_start=1
                     /transl_except=(pos:719. .721,aa:Tyr)
                     /product="75KDa protein"
                     /protein_id="AAC18569.1"
                     /db_xref="GI:3172384"
                     /translation="MVDEGRYMTWKDATHNKLMTDRWARVSSVQATINQALSLDLSKA
                     ASLPVIKANFSSLGANWSEDGDVFVAPMVRFPVTLNQFGALVLWLNLSDPAFAILMTK
                     VFTLTNAGGADNASASSRREVVAGSRPETSGKSVGTDESGSYTLTTTLQAFNQAIRLP
                     EVLWTRDKFEREWNLPWTPTRGVTPAVPRTNYQLNAARTMAHIRAALKARLYPGDSIE
                     WVGWVHSYPPPPYDGYDVPSLDIINEKLAADDVGGLVLPTPDIPGGPNFEVSEDIEQT
                     GRNGLWLMVGLLLAALAVGVGTAAYHRRKLRSRLVELKSLWFSRGSGGGGDFTVELPT
                     RTTDAFSLGTTISEHAPPPVGTMRHRSTTATDSHEALPFETWVYLNLSRVYFSIGCSD
                     LFYATREFVTVFNGEFEGLIELLEASDEDDGVYTDAESAVIVGTSAIVNDHDQFDLNT
                     VILEKRIKFKRLSLEEAELNRQESVTIEFADEERRTLMHKLESDRLEATHKVNKAAAD
                     AEAALSVAVLAAKEARTYDDKLAFDRACKEKELRLRELEVERMPGKTERYIHTGIQGG
                     AQLAGALAVGSMLRTRGVSNITPSNASGSPTGIGRVRSASQPVPTSLSGGSVLGANNN
                     NSVGSVSSGSRVESLGYNSAIRNHNTFWSKLRIGSEWRQNSISRMQARPLAGGGIPPS
                     APAV"
     CDS             146. .721
                     /note="21KDa"
                     /codon_start=1
                     /product="coat protein"
                     /protein_id="AAC18568.1"
                     /db_xref="GI:3136265"
                     /translation="MVDEGRYMTWKDATHNKLMTDRWARVSSVQATINQALSLDLSKA
                     ASLPVIKANFSSLGANWSEDGDVFVAPMVRFPVTLNQFGALVLWLNLSDPAFAILMTK
                     VFTLTNAGGADNASASSRREVVAGSRPETSGKSVGTDESGSYTLTTTLQAFNQAIRLP
                     EVLWTRDKFEREWNLPWTPTRGVTPAVPRTN"
     CDS             2125. .3273
                     /codon_start=1
                     /product="42KDa protein"
                     /protein_id="AAC18570.1"
                     /db_xref="GI:3136267"
                     /translation="MAPEQHKQNASETASGRRNSSVRSRGMSKDDWSVTHPDDVFSII
                     EKTLVEDGYKWHGVKPGHCDWDKLEQSGAIKNFKGTLEGEVDSSCSLTCNAAAIKLDI
                     VERLDVSSDWSARVGIVLGAPGVGKSTSIKHILDTYGSRYKMVLCLPVKQLLDGVFSG
                     RMDTFLIDDIFSRSVDYGKYHTMLVDEITRVHMCEVLVLAGYLGIKNVICFGDPAQGI
                     NFKAGSAVNYNFPVIAECYSSRRFGVATADLINSCNGGGKSVVGNNDVKDNWTFEELC
                     GKIEEMSTVLVATHATKEFLADDGIEAVYYEDAQGMTYDVVTIVLKDEFDDDAICDSN
                     VRAVLLTRARKGGLLKVDPNIAARFKNGDFNSRGVSKACTGDTFCEDR"
     CDS             3273. .3629
                     /codon_start=1
                     /product="13KDa protein"
                     /protein_id="AAC18571.1"
                     /db_xref="GI:3136268"
                     /translation="MSREITARANKNVPIVVGVCVVAFFVLLAFMQQKHKTHSGGDYG
                     VPTFSNGGKYRDGTRSADFNSNNHRAYGCGGSKSSVTGKVGQQLLVLALVVAVFVLFM
                     RGCWSSPEHICNGSCG"
     CDS             3307. .3513
                     /note="putative open reading frame nested within the 13
                     kDa gene"
                     /codon_start=1
                     /product="8KDa protein"
                     /protein_id="AAC18572.1"
                     /db_xref="GI:3136269"
                     /translation="MCLLLLVFVLLLSLYCWLSCNKNIRRTLVAITGFQLFRTVESIG
                     MGLGLRTLIVIIIVLMAVEVLNLV"
     CDS             3613. .4011
                     /codon_start=1
                     /product="15KDa protein"
                     /protein_id="AAC18573.1"
                     /db_xref="GI:3136270"
                     /translation="MVLVVKVDFSTIVLYIVAGVVVVSVLYSPFFSNEVKAGGYAGAI
                     FPNGGCIMDRNSFAQFGGCDIPKYVADSISRVAIKELDADIKADLNSVVAKRVVLYEG
                     LAQLCYRVFSWLVCLFMVCLMLFVWFWYHS"
     CDS             4039. .4401
                     /codon_start=1
                     /product="14KDa protein"
                     /protein_id="AAC18574.1"
                     /db_xref="GI:3136271"
                     /translation="MEKSNSIGVYVKDPITNDCRLFSVKCGNWCLFTNHVFVTYRGKN
                     DDEKVVKDTCRLHFHVKCVSCSSKVTFKANNRDHLEWLSKGFVRVNRNFSIVGACSKC
                     RGVFDSCAQQDELDNNVV"
     polyA_site      4616
BASE COUNT     1124 a    779 c   1220 g   1493 t
ORIGIN
```

```
  Query Match             58.3%;  Score 208.2;  DB 14;  Length 4616;
  Best Local Similarity   73.9%;  Pred. No. 6.8e-51;
  Matches  264;  Conservative    0;  Mismatches   93;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
           |||||||| |||||||| || ||  | ||||||||||||||||||||||||||||||||
Db    3273 ATGTCTAGAGAAATAACGGCCCGTGCTAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 3332

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGG 120
           ||||| |||||||||||||||||||| |||||||| |||  |  |  ||| || |||||
Db    3333 GTTGTTGCTTTCTTTGTATTGCTGGCTTTCATGCAACAAAAACATAAGACGCACTCTGGT 3392

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
           || |||||||| || ||||| ||||| |||||||| |  ||||| || || || || ||
Db    3393 GGCGATTACGGGGTTCCAACTTTTTCGAACGGTGGAAAGTATAGGGATGGGACTAGGTCT 3452

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 240
           || || |||||||||||| |||||||||||||| || || || || |||     ||| ||
Db    3453 GCGGACTTTAATAGTAATAATCATCGTGCTTATGGCTGTGGAGGTTCTAAATCTAGTGTG 3512

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
           | | |    |||||||||||| ||  ||||| | ||  | ||    |   |  | || |
Db    3513 ACTGGCAAGGTTGGGCAGCAGCTGCTTGTGCTTGCCCTAGTAGTAGCTGTCTTTGTTTTG 3572

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
           |  | |   | |   ||||  | || || || ||||| |||||| ||||||||||
Db    3573 TTCATGCGTGGTTGCTGGTCCTCTCCTGAGCATATTTGCAATGGTTCTTGTGGTTAA 3629


RESULT 13
LRVBABCD
LOCUS       LRVBABCD                3065 bp    RNA     linear   VRL 06-DEC-1995
DEFINITION  Lychnis ringspot virus RNA for beta-A, beta-B, beta-C, beta-D
            proteins.
ACCESSION   Z46351
VERSION     Z46351.1  GI:1107714
KEYWORDS    coat protein.
SOURCE      Lychnis ringspot virus.
  ORGANISM  Lychnis ringspot virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Hordeivirus.
REFERENCE   1  (bases 1 to 3065)
  AUTHORS   Solovyev,A.G.
  JOURNAL   Unpublished
```

```
  REMARK     (sites)
REFERENCE    2  (bases 1 to 3065)
  AUTHORS    Solovyev,A.
  TITLE      Direct Submission
  JOURNAL    Submitted (21-OCT-1994) Solovyev A., Institute of Biochemistry and
             Plant Virology, Messeweg 11/12, Braunschweig, Germany
FEATURES             Location/Qualifiers
     source          1. .3065
                     /organism="Lychnis ringspot virus"
                     /db_xref="taxon:44421"
     misc_feature    1. .3065
                     /note="genomic RNA beta, partial sequence"
     CDS             157. .744
                     /function="coat protein"
                     /citation=[1]
                     /codon_start=1
                     /product="beta-A protein"
                     /protein_id="CAA86469.1"
                     /db_xref="GI:1107715"
                     /db_xref="SPTREMBL:Q83076"
                     /translation="MANLGLTVKGGGHYDEDQWNSNIIDSKVIYDWWISLDSWHSLIS
                     TLSKVQTDVVSSRRAIADLIDGVKKDLPAALDRRFAGSQGVLGAANYRAPPLFVRLDS
                     AMRANLIAIARMGDQPFLSRDQELSRNRPNTGTNQQPSNALVARDVQPLRDSALHFTY
                     DLKDLVTSDPPVFDRRSFEAKFQLNWVARAPGDNA"
     CDS             850. .2211
                     /function="virus cell-to-cell movement protein"
                     /citation=[1]
                     /codon_start=1
                     /product="beta-B protein"
                     /protein_id="CAA86470.1"
                     /db_xref="GI:1107716"
                     /db_xref="SPTREMBL:Q83077"
                     /translation="MNIFLVYNTMAIDKIVKSDNRSQKQQSSKKKKNRKHGNKVVENS
                     TNVGVPPNEGVTLRRVRSVESIDEIFTSNQRTPPAAEQLPQKQKQSKPPSVAKKHAAN
                     DKRRYEKADAFMNKIVEASRASGFQPTGQNFKRCPANLFEKCKLRALYDKHLRVLHKD
                     ACDSEKERIAAKSFLHRSLRPNVDFKVGIVSGVAGSGKSTLIRKLCSEADAMCVLANP
                     RLKETDYKGQSKTFTLQQVLLSIVPMTSDIVIVDEYTLAESAELLLLQRKLQATFLVL
                     FGDVAQGNAKTASSLEYLQFPVVFISKTSHRLGKHTAELCKKHGQAFEPGSPEEDEII
                     VADYLGAADTTEKNIAFTKETVEDLRDAGVEASLVLETQGKEYESVTLFIRESDEAAM
                     ADSHLRAVALTRHRKKLIIRAEPGVQSSFLNGELKSKTSADSHKYESSKVSYADSSSA
                     AAQ"
     CDS             2180. .2551
                     /function="virus cell-to-cell movement protein"
                     /citation=[1]
                     /codon_start=1
                     /product="beta-D protein"
                     /protein_id="CAA86471.1"
                     /db_xref="GI:1107717"
                     /db_xref="SPTREMBL:Q83078"
                     /translation="MPTQVAQRPNKYWPIVVGVGLIGLFAYLIFTNQKHATQSGDNIH
                     KFANGGSYQDGNKRINYNKNNNLAYGYKGLSNASSVDLWMLGLCIAAIAAGIYGEYLR
                     KRKHDECTTCPPDCKICGGWR"
     CDS             2406. .2879
                     /function="virus cell-to-cell movement protein"
                     /citation=[1]
                     /codon_start=1
                     /product="beta-C protein"
                     /protein_id="CAA86472.1"
                     /db_xref="GI:1107718"
                     /db_xref="SPTREMBL:Q83079"
                     /translation="MPHPLTCGCSDCVLPPLQREFTESIYANENMMSAPPAPQTAKSV
                     VGGDDLFLMMYSFLAGVLLTLLVIWSVGGTTCSNPQKASYYYQDLNKIEMEVAPGSPI
                     DPEVIKAIHHFQNFPYGRTPGLGWIDDVTTLVTVWFTRLIYLGIIFFFIWVFKNI"
BASE COUNT      889 a    597 c    681 g    898 t
ORIGIN

  Query Match             18.3%;  Score 65.4;  DB 14;  Length 3065;
  Best Local Similarity   59.9%;  Pred. No. 1.9e-08;
  Matches  130;  Conservative    0;  Mismatches   81;  Indels    6;  Gaps    1;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
             |||  |  |   ||  ||  |    ||  |||||||||  ||  ||||||  ||||||||||  ||
Db    2180 ATGCCGACTCAAGTAGCGCAGCGGCCCAATAAGTATTGGCCTATAGTTGTTGGTGTGGGT 2239

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGG 120
           |  |  |  |  ||||  |  |||   ||  |   | |||      ||  ||  ||  ||  ||
Db    2240 TTGATAGGTCTTTTTGCTTATCTGATCTTTACCAATCAAAAGCACGCAACGCAATCAGGT 2299

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
           |  |  ||       |  |  | |||| | |||||||||   |||  ||||||||  |  |  |
Db    2300 GATAATA------TTCACAAATTTGCCAACGGTGGTTCTTATCAAGACGGTAATAAAAGA 2353

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGT 217
               ||| ||| |  ||  |||  ||   |||||||| |
Db    2354 ATAAATTATAACAAAAATAATAATTTAGCTTACGGTT 2390


RESULT 14
BSBV3GNS
LOCUS       BSBV3GNS                3005 bp    RNA     linear   VRL 05-DEC-2000
DEFINITION  Beet soil-borne virus genes for 13K, 22K and 48K proteins.
ACCESSION   Z66493
VERSION     Z66493.2  GI:11595427
KEYWORDS    13K protein; 22K protein; 48K protein.
SOURCE      Beet soil-borne virus.
  ORGANISM  Beet soil-borne virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Pomovirus.
REFERENCE   1  (bases 1 to 3005)
  AUTHORS   Koenig,R., Beier,C., Commandeur,U., Luth,U., Kaufmann,A. and
            Luddecke,P.
  TITLE     Beet soil-borne virus RNA 3--a further example of the heterogeneity
            of the gene content of furovirus genomes and of triple gene
            block-carrying RNAs
  JOURNAL   Virology 216 (1), 202-207 (1996)
  MEDLINE   96187804
   PUBMED   8614988
REFERENCE   2  (bases 1 to 3005)
  AUTHORS   Koenig,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-1995) Koenig R., Biologische Bundesanstalt fuer
            Land- und Forstwirtschaft, Institut fuer
            Biochemie/Pflanzenvirologie, Messeweg 11, Braunschweig, Germany,
            D38104
  REMARK    Revised by [3]
REFERENCE   3  (bases 1 to 3005)
  AUTHORS   Koenig,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-DEC-2000) Koenig R., Biologische Bundesanstalt fuer
            Land- und Forstwirtschaft, Institut fuer
            Biochemie/Pflanzenvirologie, Messeweg 11, Braunschweig, Germany,
            D38104
COMMENT     On Dec 6, 2000 this sequence version replaced gi:1225930.
FEATURES             Location/Qualifiers
     source          1. .3005
                     /organism="Beet soil-borne virus"
                     /strain="ahlum"
                     /db_xref="taxon:46436"
     5'UTR           1. .404
     CDS             405. .1688
                     /function="Triple gene block protein, presumably cell to
                     cell movement"
                     /note="putative. NTP-binding motif present on amino acids
                     179-186"
                     /citation=[1]
                     /codon_start=1
                     /product="48K protein"
                     /protein_id="CAA91254.1"
                     /db_xref="GI:1225931"
                     /db_xref="SPTREMBL:Q65726"
                     /translation="MEKEKLQKKERNLNRRTNKKGVRRLDKKNSEEKTKEDNRKSVEE
                     DVVDGLGAHESVAKDRAATCERGKSQRQRVDQETSVKEGAGVDSKLGSDRYAGKRQLE
                     VVSRICQESGFVATGKPLKRYPEDYFLKSGLLADFDKYLSDRLDKGCNLTKSETETVL
                     KHLRQKRRPQSFLAGSVTGVPGSGKTTLLRKIQTEAGLNSVVILANERHKIRFTQLPA
                     CYTAKEILLLRTAIKYDVLLIDEYTLLQNGEILLLQRILEAKVVVLFGDRAQGNSRTA
                     DSPEWLQIPVIYSSVKSRRFGKATADFCGKQGFDFEGCDQEDEVQKLDFEGSSPETDI
                     NLALTEATIEDLKEVGIECSLVKDVQGNEYDSVSLFIREEDRAALSDPELRSVAFTRH
```

```
                RKLLIVRIPVCLMLSLFNGELNSDYRPQTNHYGKN"
     CDS        1675. .2025
                /function="Triple gene block protein, presumably cell to
                cell movement"
                /note="putative"
                /citation=[1]
                /codon_start=1
                /product="13K protein"
                /protein_id="CAA91255.1"
                /db_xref="GI:1225932"
                /db_xref="SPTREMBL:Q65727"
                /translation="MVRTNEIGARPNKYWPIVVGVFAIALFSFLTITNQKHATESGDN
                IHKFSNGGKFQDGNKRVHYNKNNPRAYNGSSSNNTFSKLLLPGLFLAAIMYAYVQYRK
                PQCTVTCRGECAHG"
     CDS        1889. .2461
                /function="triple gene block protein, presumable cell to
                cell movement"
                /note="putative"
                /citation=[1]
                /codon_start=1
                /product="22K protein"
                /protein_id="CAA91256.1"
                /db_xref="GI:1225933"
                /db_xref="SPTREMBL:Q65728"
                /translation="MDPPAIIHSQNCCCPDCSWQPSCTHTSNTESLNARSPVEENVRM
                VKGYSTDYVIMCVIVSVSLGFAIAAYFYSSGGHYDPDAAIFKQDLNEVQINFGKHPID
                PKVIEAVHHWQRAPFGRFQGESVDVSKNFFEDSVDEDDEEERAEVNKEEVTKKKSKVA
                VCNRLCSIRNSLLRLFWVITLIIIVVKCRF"
     3'UTR      2462. .3005
BASE COUNT      917 a    507 c    664 g    917 t
ORIGIN

  Query Match           16.7%;  Score 59.6;  DB 14;  Length 3005;
  Best Local Similarity  59.1%;  Pred. No. 9.9e-07;
  Matches  123;  Conservative    0;  Mismatches   79;  Indels    6;  Gaps    1;

Qy      5 CTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTG 64
          |||  || ||   |||  |||| |||||| ||  ||| || || || |||||  | ||
Db   1685 CTAATGAGATCGGCGCCAGACCTAATAAGTATTGGCCAATAGTCGTGGGTGTATTTGCAA 1744

Qy     65 TGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGGGGTG 124
          | || || ||    ||  |  |  | |   | ||   |   || ||| | || || | |
Db   1745 TTGCATTGTTCAGTTTTTTAACTATAACTAATCAGAAACACGCCACAGAGTCGGGAGATA 1804

Qy    125 ATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGCTG 184
          ||       |||  |  |||||||||||||| | || |  |||||| |  |      | |
Db   1805 ATA------TCCACAAGTTTTCTAACGGTGGAAAATTTCAAGACGGCAATAAGCGTGTTC 1858

Qy    185 ATTTTAATAGTAACAATCATCGTGCTTA 212
          ||| |||||  || |||| | | |||||
Db   1859 ATTATAATAAGAATAATCCTAGAGCTTA 1886


RESULT 15
PSZBETACD
LOCUS       PSZBETACD               3591 bp ss-RNA     linear   VRL 02-AUG-1996
DEFINITION  Poa semilatent virus beta genomic segment beta-A, beta-B, beta-C
            and beta-D protein genes, complete cds.
ACCESSION   M81486
VERSION     M81486.1  GI:1478094
KEYWORDS    .
SOURCE      Poa semilatent virus cDNA to genomic RNA.
  ORGANISM  Poa semilatent virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Hordeivirus.
REFERENCE   1  (bases 1 to 3591)
  AUTHORS   Agranovsky,A.A., Karasev,A.V., Novikov,V.K., Lunina,N.A.,
            Loginov,S. and Tyulkina,L.G.
  TITLE     Poa semilatent virus, a hordeivirus having no internal polydisperse
            poly(A) in the 3' non-coding region of the RNA genome
  JOURNAL   J. Gen. Virol. 73 (Pt 8), 2085-2092 (1992)
  MEDLINE   92356087
   PUBMED   1645144
REFERENCE   2  (bases 1 to 3591)
  AUTHORS   Solovyev,A.G., Savenkov,E.I., Agranovsky,A.A. and Morozov,S.Y.
  TITLE     Comparisons of the genomic cis-elements and coding regions in RNA
            beta components of the hordeiviruses barley stripe mosaic virus,
            lychnis ringspot virus, and poa semilatent virus
  JOURNAL   Virology 219 (1), 9-18 (1996)
  MEDLINE   96204567
   PUBMED   8623558
REFERENCE   3  (bases 1 to 3591)
  AUTHORS   Solovyev,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-AUG-1993) Andrey G. Solovyev, Institute of
            Biochemistry and Plant Virology, Braunschweig, D-38104, Germany
COMMENT     On Aug 2, 1996 this sequence version replaced gi:333326.
FEATURES             Location/Qualifiers
     source          1. .3591
                     /organism="Poa semilatent virus"
                     /db_xref="taxon:12328"
                     /chromosome="RNA beta"
     CDS             97. .699
                     /function="coat protein"
                     /codon_start=1
                     /product="beta-A protein"
                     /protein_id="AAB05576.1"
                     /db_xref="GI:1478095"
                     /translation="MPNISLTTKGGGHYDEDLWFSQVVSQGLYAEWWVLSDNWEKFLD
                     NLRGVNFGVSSSRSQVSDYVAALDRDLPADVDRRFAGPRGQLGSPNYVSAPIFYRINR
                     SLVGDITRLARITDQPSNNRDVELSRVKKLSGQPTPPPGTDNQMAIRDVQSLRDGSLS
                     FSYTLINELVGAQLAAYTVRRFETEFGLQWLPVQAQGQRD"
     CDS             826. .2556
                     /function="virus cell-to-cell movement protein"
                     /codon_start=1
                     /product="beta-B-protein"
                     /protein_id="AAB05577.1"
                     /db_xref="GI:1478096"
                     /translation="MSDGLRTNHQNETPGVQSEKMELRQTTSDDTQNQADSGAKSHDT
                     NTSRVDETAKTTDVKESPSGESLSGVSAVSGGTDKGEKSAVRKEAGESKESPKSVPSD
                     GGEVKHVDTETKAKSKRKKKNKKTPKEGTSKTTSESSSAKNVESKESKKQTKPKAVSP
                     SSDTSKIASEVGSAKKATKKESKKQTKDKGSAEDLNANTKLKAKASEQKGPTIPGTSA
                     EASRIDLLQSSTAEKFTKNDVRRTALVNEFVAQIHKFCIEQGFEPTGRQYMRARANLF
                     ELVGLRNLYMEHLKKTAAKACNFTKDRIRRKLFLTSNHKPSVDFLVGIVSGVPGCGKS
                     TLVRKLLDSPISCYVALANPATERDYRGTSNVMTLDDLLLAKVPMSSDLLIVDEYTLA
                     ESAEILLLQRRVRSSLVLLVGDVAQGRSNNASNIEYLTLPVIYRKITSHRIGEETAKA
                     CSKQGNRIRSAGRKDKLILVDYEGETEETEKNLAFTEETRDDVKDCGYDCSLVSEVQG
                     LEYESVTLFLRNTDRAAASDHHKRTVSMTRHKSLLIIRAEQEIGQPFLMGELSVSSKR
                     PSNAHVYSSE"
     CDS             2559. .2969
                     /function="virus cell-to-cell movement protein"
                     /codon_start=1
                     /product="beta-D protein"
                     /protein_id="AAB05578.1"
                     /db_xref="GI:1478097"
                     /translation="MPAKTTSVGSRPNKYWPIVAGIGVVGLFAYLIFANQKHSTESGD
                     NIHKFANGGSYRDGSKCISYHRNNPFAYGNASSPGMLLPTLFTIIGIVSYLWATRGEI
                     MGGNSPLLGHNCGEECAGECVRGLDRQELPPNSV"
     CDS             2773. .3261
                     /function="virus cell-to-cell movement protein"
                     /codon_start=1
                     /product="beta-C protein"
                     /protein_id="AAB05579.1"
                     /db_xref="GI:1478098"
                     /translation="MAMPHLLGCSCPLCSQSSESFPIYGPPEERLWEETPPSSATIVE
                     RNAPENVFVAWIDRNYLLILCSLALLSVSLAYIYFTSGNNNPYVKGGYFYQDLNSVEV
                     RFGEHPVDPKVIASIHHWQKNPFGVSPLWEGLNNLVSSLVSALKFYVGLLFLLLLIII
                     FK"
BASE COUNT      1049 a    718 c    842 g    982 t
ORIGIN

  Query Match           14.8%;  Score 53;  DB 14;  Length 3591;
  Best Local Similarity  58.5%;  Pred. No. 9.1e-05;
  Matches  114;  Conservative    0;  Mismatches   75;  Indels    6;  Gaps    1;

Qy     21 TCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATT 80
          | |||| |||||| ||  ||| ||||| | |||| |  |||| || | ||| ||||  |
```

```
Db   2588 TAGACCTAATAAGTATTGGCCAATTGTAGCTGGTATAGGTGTAGTAGGTTTATTTGCTTA 2647

Qy     81 GCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGGGGGTGATTACGGAGTCCCAAC 140
          |    ||      | ||   |  | | ||  | |||| | | ||       |||  |
Db   2648 CTTAATTTTTGCTAATCAGAAACATTCTACTGAGTCTGGAGATAATA------TCCACAA 2701

Qy    141 ATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAA 200
          |||| | |||||||| |  ||| |||||||| | |  |     |  ||   || | |||||
Db   2702 ATTTGCCAACGGTGGGAGTTATCGAGACGGTTCTAAGTGTATTTCTTACCATCGAAACAA 2761

Qy    201 TCATCGTGCTTACGG 215
          | |   || || ||
Db   2762 CCCTTTCGCCTATGG 2776


Search completed: March 12, 2003, 09:31:48
Job time : 1177 secs
```

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: March 11, 2003, 14:24:19 ; Search time 175 Seconds
(without alignments)
4594.077 Million cell updates/sec

Title: US-09-936-011-3
Perfect score: 357
Sequence: 1 atgtctagggaaataaccgc.....gtaatggtgcttgtggttaa 357

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	357	100.0	357	21	AAA75878	DNA encoding a P13
	2	347.4	97.3	357	21	AAA75877	DNA encoding a P13
	3	328.2	91.9	739	22	AAF26890	Beet necrotic yell
	4	328.2	91.9	4609	22	AAF26889	Beet necrotic yell
	5	315.4	88.3	739	22	AAF26891	Beet necrotic yell
	6	50.6	14.2	507	24	ABK86954	Potato mop-top vir
	7	42.6	11.9	507	24	ABK86955	Mutant potato mop-
c	8	36.8	10.3	1230025	20	AAX91990	Nucleotide sequenc
	9	36.6	10.3	4039	22	AAH34822	Human colon cancer
	10	36.6	10.3	4039	24	ABL90346	Human polynucleoti
	11	36.6	10.3	4040	21	AAC77977	Human cancer assoc
	12	36.6	10.3	5544	24	ABL61761	Colon adenocarcino
	13	36.6	10.3	5549	23	AAS87055	DNA encoding novel
c	14	36.4	10.2	2215	24	ABI99688	Mouse ischaemic co
	15	35.6	10.0	1668	24	ABK40083	Human chemically p
	16	35.2	9.9	14920	24	ABN80146	Human chemically m
c	17	34.8	9.7	2231	19	AAV32556	Candida albicans C
c	18	34.8	9.7	3243	19	AAV32555	Candida albicans C
c	19	34.8	9.7	3297	23	AAS91437	DNA encoding novel
	20	34.6	9.7	2772	23	ABL25868	Drosophila melanog
c	21	34.2	9.6	493	22	ABA58243	Human foetal liver
c	22	34.2	9.6	493	22	ABA27397	Probe #5863 for ge
c	23	34.2	9.6	493	22	AAK06331	Human brain expres
c	24	34.2	9.6	493	22	AAK31999	Human bone marrow
c	25	34.2	9.6	493	22	AAI37860	Probe #6546 used t
c	26	34.2	9.6	493	24	ABS06764	Human genome-deriv
c	27	34.2	9.6	2328	23	AAS84696	DNA encoding novel
	28	34.2	9.6	2494	24	ABK31539	Signal transductio
c	29	34.2	9.6	5296	22	AAK77661	Human immune/haema
c	30	34.2	9.6	47670	23	ABL16824	Drosophila melanog
c	31	34.2	9.6	50368	23	ABL16768	Drosophila melanog
c	32	33.8	9.5	406	24	ABN75893	Human ORF840 cDNA,
c	33	33.8	9.5	1794	23	AAS78187	DNA encoding novel
c	34	33.8	9.5	3598	22	AAH18642	Human cDNA sequenc
	35	33.2	9.3	5971	24	ABL32383	Human immune syste
c	36	32.8	9.2	2361	23	AAS54928	Staphylococcus aur
	37	32.6	9.1	5986	24	AAS61433	Human gene regulat
	38	32.6	9.1	5986	24	ABK31499	Signal transductio
	39	32.6	9.1	6681	24	ABL54303	Chemically treated
	40	32.6	9.1	6681	24	ABL32154	Human immune syste
c	41	32.4	9.1	2527	15	AAQ56756	Chitinase 1 gene.
	42	32.4	9.1	7434	24	AAD28386	Human chemically t
c	43	32.4	9.1	29376	23	ABL08834	Drosophila melanog
c	44	32.4	9.1	107602	24	AAK99657	DNA of the PAC clo
c	45	32.4	9.1	107612	24	ABL54503	Human PAC clone se

ALIGNMENTS

RESULT 1
AAA75878
ID AAA75878 standard; DNA; 357 BP.
XX
AC AAA75878;
XX
DT 22-JAN-2001 (first entry)
XX
DE DNA encoding a P13 variant of Beet necrotic yellow vein mosaic virus.
XX
KW P13 protein; triple gene block 2; TGB2; resistance; group I virus; ss.
XX
OS Beet necrotic yellow vein mosaic virus.
XX
FH Key Location/Qualifiers
FT CDS 1..357
FT /*tag= a
FT /product= "P13 protein"
XX
PN WO200055301-A2.
XX
PD 21-SEP-2000.
XX
PF 07-MAR-2000; 2000WO-EP02176.
XX
PR 12-MAR-1999; 99EP-0200773.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Jonard G, Lauber E, Guilley H, Richards K;
XX
DR WPI; 2000-602114/57.

```
DR   P-PSDB; AAB18802.
XX
PT   Inducing viral resistance into a plant or a plant cell, especially beet
PT   necrotic yellow vein virus-resistance into a sugar beet plant or cell
PT   by transforming the plant cell with triple gene block of the virus  -
XX
PS   Disclosure; Fig 2; 27pp; English.
XX
CC   The present sequence encodes a P13 protein variant of Beet necrotic
CC   yellow vein mosaic virus. The P13 sequence is a triple gene block 2
CC   (TGB2) sequence. It is used for inducing resistance to a group I virus.
CC   The method is useful for inducing resistance to a plant cell or plants
CC   such as sugar beet, potato, barley or peanut against group I virus
CC   such as hordeiviruses, benyviruses, pecluviruses and pomoviruses,
CC   preferably barley stripe mosaic virus, potato mop top virus, peanut
CC   clump virus and the beet soil-borne virus, more preferably beet
CC   necrotic yellow vein virus (BNYVV).
XX
SQ   Sequence 357 BP; 84 A; 60 C; 96 G; 117 T; 0 other;

  Query Match             100.0%;  Score 357;  DB 21;  Length 357;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-101;
  Matches  357;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGG 120

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357


RESULT 2
AAA75877
ID   AAA75877 standard; DNA; 357 BP.
XX
AC   AAA75877;
XX
DT   22-JAN-2001  (first entry)
XX
DE   DNA encoding a P13 protein of Beet necrotic yellow vein mosaic virus.
XX
KW   P13 protein; triple gene block 2; TGB2; resistance; group I virus; ss.
XX
OS   Beet necrotic yellow vein mosaic virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..357
FT                   /*tag=  a
FT                   /product= "P13 protein"
XX
PN   WO200055301-A2.
XX
PD   21-SEP-2000.
XX
PF   07-MAR-2000; 2000WO-EP02176.
XX
PR   12-MAR-1999;   99EP-0200773.
XX
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Jonard G,  Lauber E,  Guilley H,  Richards K;
XX
DR   WPI; 2000-602114/57.
DR   P-PSDB; AAB18801.
XX
PT   Inducing viral resistance into a plant or a plant cell, especially beet
PT   necrotic yellow vein virus-resistance into a sugar beet plant or cell
PT   by transforming the plant cell with triple gene block of the virus  -
XX
PS   Disclosure; Fig 1; 27pp; English.
XX
CC   The present sequence encodes a P13 protein of Beet necrotic yellow
CC   vein mosaic virus. The sequence is a triple gene block 2 (TGB2)
CC   sequence. It is used for inducing resistance to a group I virus. The
CC   method is useful for inducing resistance to a plant cell or plants
CC   such as sugar beet, potato, barley or peanut against group I virus
CC   such as hordeiviruses, benyviruses, pecluviruses and pomoviruses,
CC   preferably barley stripe mosaic virus, potato mop top virus, peanut
CC   clump virus and the beet soil-borne virus, more preferably beet
CC   necrotic yellow vein virus (BNYVV).
XX
SQ   Sequence 357 BP; 89 A; 58 C; 93 G; 117 T; 0 other;

  Query Match             97.3%;  Score 347.4;  DB 21;  Length 357;
  Best Local Similarity   98.3%;  Pred. No. 5.5e-98;
  Matches  351;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGG 120
           |||||||||||||||||||||||||||||||||||  |   |  ||||||||||||||
Db      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357


RESULT 3
AAF26890
ID   AAF26890 standard; DNA; 739 BP.
XX
AC   AAF26890;
XX
DT   09-APR-2001  (first entry)
XX
DE   Beet necrotic yellow vein virus promoter sequence SEQ ID NO:2.
XX
KW   Beet necrotic yellow vein virus; BNYVV; transformed plant;
KW   Rhizomania disease-resistant plant; promoter; ds.
XX
OS   Beet necrotic yellow vein mosaic virus.
XX
PN   JP2000312540-A.
```

```
XX
PD   14-NOV-2000.
XX
PF   28-APR-1999;   99JP-0122628.
XX
PR   28-APR-1999;   99JP-0122628.
XX
PA   (HOKK-) HOKKAIDO PREFECTURE.
PA   (HOKK-) HOKKAIDO TENSAI KYOKAI SH.
XX
DR   WPI; 2001-054202/07.
XX
PT   A Rhizomania disease-resistant plant  -
XX
PS   Claim 14; Page 9; 11pp; Japanese.
XX
CC   The present invention describes a method for producing a transformed
CC   plant in which resistance against beet necrotic yellow vein virus
CC   (BNYVV) is given by transforming expressably a gene derived from BNYVV
CC   genome or a DNA corresponding to its part or a DNA substantially same
CC   as it in a plant genome. The vector structure can be used for
CC   transforming a plant or a plant cell having BNYVV resistance. The
CC   present sequence represents a wild type BNYVV promoter nucleotide
CC   sequence for use in the method of the invention.
XX
SQ   Sequence 739 BP; 184 A; 114 C; 189 G; 252 T; 0 other;

  Query Match             91.9%;  Score 328.2;  DB 22;  Length 739;
  Best Local Similarity   95.0%;  Pred. No. 6.9e-92;
  Matches  339;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGG 120
          |||||||||||||||||||||||||||||||||||||||  |  |  ||||||||||||||
Db      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
           || ||||||||||||||||||||||||||||||||| ||| ||||||||||||||| |||
Db .   121 GGCGATTACGGAGTCCCAACATTTTCTAACGGTGGTAAATACAGAGACGGTACAAGGTCA 180

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db     181 GCTGATTTTAATAGTAATAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
           |||||||||||| |||||||||||| |||||||||||||||| ||||||||||||||| |||
Db     241 AGTAGTCGAGTCGGGCAGCAACTTGTTGTGTTAGCTATTGTGTCTGTGTTAATAGTATCA 300

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
           || |||||||||||||| |||||||||||||| |||||||||||||||||||||||||
Db     301 CTGTTACAACGATTAAGATCTCCACCAGAACATATTTGTAATGGTGCTTGTGGTTAA 357


RESULT 4
AAF26889
ID   AAF26889 standard; DNA; 4609 BP.
XX
AC   AAF26889;
XX
DT   09-APR-2001  (first entry)
XX
DE   Beet necrotic yellow vein virus RNA-2 nucleotide sequence SEQ ID NO:1.
XX
KW   Beet necrotic yellow vein virus; BNYVV; transformed plant;
KW   Rhizomania disease-resistant plant; ds.
XX
OS   Beet necrotic yellow vein mosaic virus.
XX
PN   JP2000312540-A.
XX
PD   14-NOV-2000.
XX
PF   28-APR-1999;   99JP-0122628.
XX
PR   28-APR-1999;   99JP-0122628.
XX
PA   (HOKK-) HOKKAIDO PREFECTURE.
PA   (HOKK-) HOKKAIDO TENSAI KYOKAI SH.
XX
DR   WPI; 2001-054202/07.
XX
PT   A Rhizomania disease-resistant plant  -
XX
PS   Claim 13; Page 7-9; 11pp; Japanese.
XX
CC   The present invention describes a method for producing a transformed
CC   plant in which resistance against beet necrotic yellow vein virus
CC   (BNYVV) is given by transforming expressably a gene derived from BNYVV
CC   genome or a DNA corresponding to its part or a DNA substantially same
CC   as it in a plant genome. The vector structure can be used for
CC   transforming a plant or a plant cell having BNYVV resistance. The
CC   present sequence represents a specifically claimed BNYVV nucleotide
CC   sequence for use in the method of the invention.
XX
SQ   Sequence 4609 BP; 1187 A; 765 C; 1221 G; 1436 T; 0 other;

  Query Match             91.9%;  Score 328.2;  DB 22;  Length 4609;
  Best Local Similarity   95.0%;  Pred. No. 1.4e-91;
  Matches  339;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps .  0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3284 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 3343

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGG 120
          |||||||||||||||||||||||||||||||||||||||  |  |  ||||||||||||||
Db    3344 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 3403

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
           || ||||||||||||||||||||||||||||||||| ||| ||||||||||||||| |||
Db    3404 GGCGATTACGGAGTCCCAACATTTTCTAACGGTGGTAAATACAGAGACGGTACAAGGTCA 3463

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 240
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    3464 GCTGATTTTAATAGTAATAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGGTAGCGTT 3523

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
           |||||||||||| |||||||||||| |||||||||||||||| ||||||||||||||| |||
Db    3524 AGTAGTCGAGTCGGGCAGCAACTTGTTGTGTTAGCTATTGTGTCTGTGTTAATAGTATCA 3583

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
           || |||||||||||||| |||||||||||||| |||||||||||||||||||||||||
Db    3584 CTGTTACAACGATTAAGATCTCCACCAGAACATATTTGTAATGGTGCTTGTGGTTAA 3640


RESULT 5
AAF26891
ID   AAF26891 standard; DNA; 739 BP.
XX
AC   AAF26891;
XX
DT   09-APR-2001  (first entry)
XX
DE   Beet necrotic yellow vein virus mutant promoter sequence SEQ ID NO:3.
XX
KW   Beet necrotic yellow vein virus; BNYVV; transformed plant;
KW   Rhizomania disease-resistant plant; promoter; ds.
XX
OS   Beet necrotic yellow vein mosaic virus.
OS   Synthetic.
XX
PN   JP2000312540-A.
XX
```

```
PD   14-NOV-2000.
XX
PF   28-APR-1999;   99JP-0122628.
XX
PR   28-APR-1999;   99JP-0122628.
XX
PA   (HOKK-) HOKKAIDO PREFECTURE.
PA   (HOKK-) HOKKAIDO TENSAI KYOKAI SH.
XX
DR   WPI; 2001-054202/07.
XX
PT   A Rhizomania disease-resistant plant -
XX
PS   Claim 15; Page 9; 11pp; Japanese.
XX
CC   The present invention describes a method for producing a transformed
CC   plant in which resistance against beet necrotic yellow vein virus
CC   (BNYVV) is given by transforming expressably a gene derived from BNYVV
CC   genome or a DNA corresponding to its part or a DNA substantially same
CC   as it in a plant genome. The vector structure can be used for
CC   transforming a plant or a plant cell having BNYVV resistance. The
CC   present sequence represents a mutant BNYVV promoter nucleotide
CC   sequence for use in the method of the invention.
XX
SQ   Sequence 739 BP; 182 A; 117 C; 187 G; 253 T; 0 other;

  Query Match             88.3%;  Score 315.4;  DB 22;  Length 739;
  Best Local Similarity   92.7%;  Pred. No. 6.5e-88;
  Matches  331;  Conservative    0;  Mismatches   26;  Indels    0;  Gaps    0;

Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60

Qy      61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGG 120
           ||||||||||||||||||||||||||||||||||||  |  |  ||||||||||||
Db     61 GTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAAAACATAAGACACATTCTGGG 120

Qy     121 GGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCA 180
           || ||||||||||||||||||||||||||| || || ||  ||||  |   |||| |||
Db     121 GGCGATTACGGAGTCCCAACATTTTCTAACGCTGCTAAATTCAGAGCTGCAGCAAGGTCA 180

Qy     181 GCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 240
           ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db     181 GCTGATTTTAATAGTAATAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTT 240

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
           |||||||||| |||||||||| |||||||||||||| |||||||||||||| |||
Db     241 AGTAGTCGAGTCGGGCAGCAACTTGTTGTGTTAGCTATTGTGTCTGTGTTAATAGTATCA 300

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTAA 357
           || |||||||||||||| |||||||||||||| ||||||||||||||||||||||||
Db     301 CTGTTACAACGATTAAGATCTCCACCAGAACATATTTGTAATGGTGCTTGTGGTTAA 357


RESULT 6
ABK86954
ID   ABK86954 standard; cDNA; 507 BP.
XX
AC   ABK86954;
XX
DT   24-SEP-2002  (first entry)
XX
DE   Potato mop-top virus (PMTV) triple-gene-block (TGB)-2 cDNA.
XX
KW   Potato; antiviral; triple-gene-block 2; TGB-2; PMTV;
KW   potato mop-top virus; Pomovirus; spraing; dwarfing; mopping;
KW   chlorotic chevron; necrotic chevron; blotching; tuber cracking;
KW   necrotic conchoidal layer; viral movement; transgenic; molecular marker;
KW   resistance; agriculture; gene; ss.
XX
OS   Potato mop-top virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             37..396
FT                   /*tag= a
FT                   /product= "TGB-2"
XX
PN   WO200250281-A1.
XX
PD   27-JUN-2002.
XX
PF   18-DEC-2001; 2001WO-SE02805.
XX
PR   21-DEC-2000; 2000SE-0004755.
XX
PA   (PLAN-) PLANT SCI SWEDEN AB.
XX
PI   Melander M,  Lee M;
XX
DR   WPI; 2002-508804/54.
DR   P-PSDB; AAU79576.
XX
PT   Nucleic acid molecule for producing plants with increased resistance
PT   against infection by potato mop-top virus comprises a structural gene
PT   encoding a triple-gene-block 2 polypeptide from the virus -
XX
PS   Disclosure; Page 31; 36pp; English.
XX
CC   The invention discloses a nucleic acid molecule comprising a promoter
CC   operably linked to a structural gene encoding a triple-gene-block 2
CC   (TGB-2) polypeptide from potato mop-top virus (PMTV) which differs from
CC   the wild-type polypeptide at a number of residues. The potato crop
CC   (Solanum tuberosum) is susceptible to PMTV which belongs to the genus
CC   Pomovirus. The main symptoms of the disease are spraing, dwarfing
CC   (mopping), chlorotic and necrotic chevrons, blotching and tuber cracking
CC   and necrotic conchoidal layers. PMTV has a tripartite RNA genome, of
CC   which RNA 2 encodes for four proteins in a triple-gene-block and all
CC   three seem to be responsible for long-distance viral movement within the
CC   plant. The mutant TGB-2 is useful for producing transgenic plants, in
CC   particular, the potato plant, with an increased resistance against
CC   infection by PMTV. Portions of the mutant TGB-2 are useful as molecular
CC   markers for detecting a transgenic plant and for detecting a food product
CC   from the transgenic plant. The advantage the mutant TGB-2 brings is that
CC   it is able to render potato plants resistant by using solely a single
CC   gene which is easy to follow and maintain during a breeding program. The
CC   transgenic plants can be grown even in fields infected by PMTV, which
CC   increases the economic value of the yield obtained from such an infected
CC   field. The sequence presented is the wild-type TGB-2 cDNA from PMTV.
XX
SQ   Sequence 507 BP; 143 A; 94 C; 106 G; 164 T; 0 other;

  Query Match             14.2%;  Score 50.6;  DB 24;  Length 507;
  Best Local Similarity   56.5%;  Pred. No. 1.1e-05;
  Matches  117;  Conservative    0;  Mismatches   84;  Indels    6;  Gaps    1;

Qy       6 TAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGT 65
           ||  |||||    || ||||| |||||  ||  |||  | |||||  |  ||   ||   |
Db      48 TAACGAAATTGGAGCGCGACCAAATAAATATTGGCCGGTAGTTGTGGCAGTAGTTGCAAT 107

Qy      66 GGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGGGGTGA 125
             | | || |  ||  |  |  | |   | |||  |   || || || || || | | |
Db     108 TTGTCTTTTCGGTTTTTTAACAGTTACCAATCAAAAACACGCTACTCAATCAGGTGATAA 167

Qy     126 TTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGA 185
           |       | |  | |||| ||||||||||  | || || ||||||  | |       | |
Db     168 TA------TACATAAATTTGCTAACGGTGGCCTGTACAGGGACGGTTCTAAGAGTATTAA 221

Qy     186 TTTTAATAGTAACAATCATCGTGCTTA 212
             | |||| |||| ||||   | |||||
Db     222 GTATAATTGTAATAATCCCAGAGCTTA 248


RESULT 7
ABK86955
ID   ABK86955 standard; cDNA; 507 BP.
```

```
XX
AC   ABK86955;
XX
DT   24-SEP-2002  (first entry)
XX
DE   Mutant potato mop-top virus (PMTV) triple-gene-block (TGB)-2 cDNA.
XX
KW   Potato; antiviral; triple-gene-block 2; TGB-2; PMTV; 13kDmut;
KW   potato mop-top virus; Pomovirus; spraing; dwarfing; mopping;
KW   chlorotic chevron; necrotic chevron; blotching; tuber cracking;
KW   necrotic conchoidal layer; viral movement; transgenic; molecular marker;
KW   resistance; agriculture; gene; ss; mutant.
XX
OS   Potato mop-top virus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             37..396
FT                   /*tag= a
FT                   /product= "TGB-2 mutant (13kDmut)"
FT   mutation        replace(188,G)
FT                   /*tag= b
FT   mutation        replace(191,G)
FT                   /*tag= c
FT   mutation        replace(193,C)
FT                   /*tag= d
FT   mutation        replace(194,T)
FT                   /*tag= e
FT   mutation        replace(195,G)
FT                   /*tag= f
FT   mutation        replace(197,A)
FT                   /*tag= g
FT   mutation        replace(204,C)
FT                   /*tag= h
FT   mutation        replace(206,G)
FT                   /*tag= i
FT   mutation        replace(208,T)
FT                   /*tag= j
XX
PN   WO200250281-A1.
XX
PD   27-JUN-2002.
XX
PF   18-DEC-2001; 2001WO-SE02805.
XX
PR   21-DEC-2000; 2000SE-0004755.
XX
PA   (PLAN-) PLANT SCI SWEDEN AB.
XX
PI   Melander M,  Lee M;
XX
DR   WPI; 2002-508804/54.
DR   P-PSDB; AAU79577.
XX
PT   Nucleic acid molecule for producing plants with increased resistance
PT   against infection by potato mop-top virus comprises a structural gene
PT   encoding a triple-gene-block 2 polypeptide from the virus -
XX
PS   Example 3; Page 32; 36pp; English.
XX
CC   The invention discloses a nucleic acid molecule comprising a promoter
CC   operably linked to a structural gene encoding a triple-gene-block 2
CC   (TGB-2) polypeptide from potato mop-top virus (PMTV) which differs from
CC   the wild-type polypeptide at a number of residues. The potato crop
CC   (Solanum tuberosum) is susceptible to PMTV which belongs to the genus
CC   Pomovirus. The main symptoms of the disease are spraing, dwarfing
CC   (mopping), chlorotic and necrotic chevrons, blotching and tuber cracking
CC   and necrotic conchoidal layers. PMTV has a tripartite RNA genome, of
CC   which RNA 2 encodes for four proteins in a triple-gene-block and all
CC   three seem to be responsible for long-distance viral movement within the
CC   plant. The mutant TGB-2 is useful for producing transgenic plants, in
CC   particular, the potato plant, with an increased resistance against
CC   infection by PMTV. Portions of the mutant TGB-2 are useful as molecular
CC   markers for detecting a transgenic plant and for detecting a food product
CC   from the transgenic plant. The advantage the mutant TGB-2 brings is that
CC   it is able to render potato plants resistant by using solely a single
CC   gene which is easy to follow and maintain during a breeding program. The
CC   transgenic plants can be grown even in fields infected by PMTV, which
CC   increases the economic value of the yield obtained from such an infected
CC   field. The sequence presented is a mutant TGB-2 (13kDmut) cDNA from PMTV.
XX
SQ   Sequence 507 BP; 145 A; 95 C; 104 G; 163 T; 0 other;

  Query Match             11.9%;  Score 42.6;  DB 24;  Length 507;
  Best Local Similarity   54.1%;  Pred. No. 0.0035;
  Matches  112;  Conservative    0;  Mismatches   89;  Indels    6;  Gaps    1;

Qy       6 TAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGT 65
           ||  |||||    || ||||| ||||| ||  |||  | ||||| |  ||   ||   |
Db      48 TAACGAAATTGGAGCGCGACCAAATAAATATTGGCCGGTAGTTGTGGCAGTAGTTGCAAT 107

Qy      66 GGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGGGGTGA 125
             | | || |  ||  |  |  | |   | ||| |   || || || || || || | | |
Db     108 TTGTCTTTTCGGTTTTTTAACAGTTACCAATCAAAAACACGCTACTCAATCAGGTGATAA 167

Qy     126 TTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGA 185
           |       | |  | |||| |||||| ||    ||  || || | | | |       | |
Db     168 TA------TACATAAATTTGCTAACGCTGCCGAATTCAGGGAAGCTGCTAAGAGTATTAA 221

Qy     186 TTTTAATAGTAACAATCATCGTGCTTA 212
             | |||| |||| ||||   | |||||
Db     222 GTATAATTGTAATAATCCCAGAGCTTA 248


RESULT 8
AAX91990/c
ID   AAX91990 standard; DNA; 1230025 BP.
XX
AC   AAX91990;
XX
DT   13-SEP-1999  (first entry)
XX
DE   Nucleotide sequence of the complete genome of Chlamydia pneumoniae.
XX
KW   Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW   sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW   vaccine; neutralising epitope; ss.
XX
OS   Chlamydia pneumoniae.
XX
PN   WO9927105-A2.
XX
PD   03-JUN-1999.
XX
PF   20-NOV-1998;   98WO-IB01890.
XX
PR   04-NOV-1998;   98US-0107078.
PR   21-NOV-1997;   97FR-0014673.
XX
PA   (GEST ) GENSET.
XX
PI   Griffais R;
XX
DR   WPI; 1999-357842/30.
XX
PT   Genome sequence of Chlamydia pneumoniae
XX
PS   Claim 1; Page 291-611; 1912pp; English.
XX
CC   The present sequence represents the complete genome of Chlamydia
CC   pneumoniae, and encodes proteins AAY34584-Y35879. C. pneumoniae causes
CC   respiratory disease such as pneumonia and bronchitis and is thought
CC   to be a contributing factor in heart disease, sarcoidosis, sinusitis,
CC   purulent otitis media, erythema nodosum or pharyngitis. The polypeptides
CC   encoded by the open reading frames of the C. pneumoniae genome (see
CC   AAY34584-Y35879) can be used in immunogenic compositions as vaccines.
```

```
CC   Vectors containing C. pneumoniae nucleotides sequences can also be
CC   used as immunogenic compositions, especially where the vector directs
CC   the expression of a neutralising epitope of C. pneumoniae.
XX
SQ   Sequence 1230025 BP; 367213 A; 249833 C; 249013 G; 363589 T; 377 other;

  Query Match             10.3%;  Score 36.8;  DB 20;  Length 1230025;
  Best Local Similarity   61.5%;  Pred. No. 5.1;
  Matches   59;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy      29 ATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGT 88
           |||||||||  ||| || ||| |   ||| ||||||   | || || | |    ||| |
Db 773874 ATAAGAATGATCCTGTTCTTGCTATGGTTCGTGTTGCTCCGTTTTTAGCAGCTTTGGTGA 773815

Qy      89 TCATGCAGCAAGCAGCTGCGACACATTCTGGGGGGTG 124
           |||||     || | | | ||     |||||||| |
Db 773814 TCATGACTTTAGGAACCGTGATGATCTCTGGGGGCG 773779


RESULT 9
AAH34822
ID   AAH34822 standard; cDNA; 4039 BP.
XX
AC   AAH34822;
XX
DT   03-SEP-2001  (first entry)
XX
DE   -Human colon cancer antigen encoding cDNA SEQ ID NO:1904.
XX
KW   Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW   colorectal carcinoma; ss.
XX
OS   -Homo sapiens.
XX
PN   WO200122920-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26524.
XX
PR   29-SEP-1999;   99US-0157137.
PR   03-NOV-1999;   99US-0163280.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Barash SC,  Birse CE,  Rosen CA;
XX
DR   WPI; 2001-235357/24.
DR   P-PSDB; AAG75417.
XX
PT   Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT   useful for preventing, diagnosing and/or treating colorectal cancers -
XX
PS   Claim 1; Page 3415-3417; 9803pp; English.
XX
CC   AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC   cancer-associated nucleic acid molecules (N) and proteins (P), where
CC   the proteins are collectively known as colon cancer antigens. The colon
CC   cancer antigens have cytostatic activity and can be used in gene
CC   therapy and vaccine production. N and P may be used in the prevention,
CC   diagnosis and treatment of diseases associated with inappropriate P
CC   expression. For example, N and P may be used to treat disorders
CC   associated with decreased expression by rectifying mutations or deletions
CC   in a patient's genome that affect the activity of P by expressing
CC   inactive proteins or to supplement the patients own production of P.
CC   Additionally, N may be used to produce the colon cancer-associated Ps,
CC   by inserting the nucleic acids into a host cell and culturing the cell
CC   to express the proteins. N and P can be used in the prevention, diagnosis
CC   and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC   and AAB77789 represent sequences used in the exemplification of the
CC   present invention.
CC   N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC   missing at time of publication, meaning no sequences are present for
CC   SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ   Sequence 4039 BP; 1174 A; 790 C; 789 G; 1283 T; 3 other;

  Query Match             10.3%;  Score 36.6;  DB 22;  Length 4039;
  Best Local Similarity   49.7%;  Pred. No. 0.59;
  Matches   93;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy     123 TGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGC 182
           |||   |  |  ||||||   | |   ||  |  |  | || | |  | | ||  ||
Db     585 TGACCCCCTGGCACCAACACCCGCCAGTACTGTGACTTCCAAAGCCAGAGCCACATGTGC 644

Qy     183 TGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAG 242
           | ||   | | | |  || ||    ||    |   || ||||  ||||||||    |
Db     645 TCATCAAACTTGCATTAAGCAGTTGGCGGGAGATGGCTGTGGAGCTGGGGGTTTAAGTGA 704

Qy     243 TAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCACT 302
           | ||    ||  ||  |  ||| |   | ||    |  | ||| | |||| |||||   |
Db     705 TGGTTCTCTTTTGCTCCCTCTTYTGAGGGTAAAGCTACTGTCTTTCTTAAGAGTGTATTT 764

Qy     303 ATTACAA 309
           ||  |||
Db     765 ATGCCAA 771


RESULT 10
ABL90346
ID   ABL90346 standard; cDNA; 4039 BP.
XX
AC   ABL90346;
XX
DT   24-MAY-2002  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 908.
XX
KW   Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW   antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW   vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW   cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; human; secreted protein; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200190304-A2.
XX
PD   29-NOV-2001.
XX
PF   18-MAY-2001; 2001WO-US16450.
XX
PR   19-MAY-2000; 2000US-205515P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Birse CE,  Rosen CA;
XX
DR   WPI; 2002-122018/16.
DR   P-PSDB; ABB89937.
XX
PT   Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT   prevention of neural, immune system, muscular, reproductive,
PT   gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT   disorders -
XX
PS   Claim 4; SEQ ID NO 908; 2081pp + Sequence Listing; English.
XX
CC   The invention relates to novel genes (ABL89449-ABL90853) and proteins
CC   (ABB89040-ABB90444) useful for preventing, treating or ameliorating
CC   medical conditions e.g. by protein or gene therapy. The genes are
CC   isolated from a range of human tissues disclosed in the specification.
CC   The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC   in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC   and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC   marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
```

```
CC   (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC   haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC   disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC   colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC   (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC   epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC   and parasitic infections.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 4039 BP; 1174 A; 790 C; 789 G; 1283 T; 3 other;

  Query Match             10.3%;  Score 36.6;  DB 24;  Length 4039;
  Best Local Similarity   49.7%;  Pred. No. 0.59;
  Matches   93;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy     123 TGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGC 182
           |||   |   | ||||||    | |    ||  |  |  | || | |  | | ||  ||
Db     585 TGACCCCCTGGCACCAACACCCGCCAGTACTGTGACTTCCAAAGCCAGAGCCACATGTGC 644

Qy     183 TGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAG 242
           | ||   | | | |  || ||    ||    |   || ||||  ||||||||     |
Db     645 TCATCAAACTTGCATTAAGCAGTTGGCGGGAGATGGCTGTGGAGCTGGGGGTTTAAGTGA 704

Qy     243 TAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCACT 302
           | ||    ||  ||  |  ||| |   | ||    |  | ||| | |||| |||||   |
Db     705 TGGTTCTCTTTTGCTCCCTCTTYTGAGGGTAAAGCTACTGTCTTTCTTAAGAGTGTATTT 764

Qy     303 ATTACAA 309
           ||  |||
Db     765 ATGCCAA 771


RESULT 11
AAC77977
ID.  AAC77977 standard; cDNA; 4040 BP.
XX
AC   AAC77977;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human cancer associated gene sequence SEQ ID NO:371.
XX
KW   Human; cancer associated gene; cancer antigen; detection; cancer;
KW   diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
KW   antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW   antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW   dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW   vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW   immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW   allergic reaction; graft versus host disease; organ rejection;
KW   haemostatic; thrombolytic; cardiovascular disorder; infection;
KW   neurological disease; drug screening; ss.
XX
OS   Homo sapiens.
XX
PN   WO200055350-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05882.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-587533/55.
DR   P-PSDB; AAB43768.
XX
PT   Novel isolated nucleic acids comprising sequences encoding peptides
PT   useful for treating or diagnosing e.g. cancer -
XX
PS   Claim 1; Page 918-920; 2352pp; English.
XX
CC   AAC77607 to AAC78448 encode the human cancer associated proteins given
CC   in AAB43398 to AAB44239. The proteins can have activities based on the
CC   tissues and cells the genes are expressed in. Example of activities
CC   include: cytostatic; proliferative; vulnerary; immunomodulator;
CC   antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC   antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC   dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC   nootropic; vasotropic; antipsoriatic and antiangiogenic. The
CC   polynucleotides and polypeptides can be used for preventing, treating or
CC   ameliorating medical conditions and diagnosing pathological conditions.
CC   Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC   the present invention may be used to treat immune disorders by activating
CC   or inhibiting the proliferation, differentiation or mobilisation of
CC   immune cells, to treat disorders of haematopoietic cells, autoimmune
CC   disorders, allergic reactions, graft versus host disease and organ
CC   rejection, modulate haemostatic or thrombolytic activity, modulate
CC   inflammation, cancers, cardiovascular disorders, neurological disease and
CC   bacterial or viral infections. The peptides, nucleotides, antibodies,
CC   agonists and antagonists may be also be used in drug screens. AAC78449 to
CC   AAC78457 and AAB44240 represent sequences used in the exemplification of
CC   the present invention.
XX
SQ   Sequence 4040 BP; 1173 A; 789 C; 790 G; 1281 T; 7 other;

  Query Match             10.3%;  Score 36.6;  DB 21;  Length 4040;
  Best Local Similarity   49.7%;  Pred. No. 0.59;
  Matches   93;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy     123 TGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGC 182
           |||   |   | ||||||    | |    ||  |  |  | || | |  | | ||  ||
Db     586 TGACCCCCTGGCACCAACACCCGCCAGTACTGTGACTTCCAAAGCCAGAGCCACATGTGC 645

Qy     183 TGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAG 242
           | ||   | | | |  || ||    ||    |   || ||||  ||||||||     |
Db     646 TCATCAAACTTGCATTAAGCAGTTGGCGGGAGATGGCTGTGGAGCTGGGGGTTTAAGTGA 705

Qy     243 TAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCACT 302
           | ||    ||  ||  |  ||| |   | ||    |  | ||| | |||| |||||   |
Db     706 TGGTTCTCTTTTGCTCCCTCTTYTGAGGGTAAAGCTACTGTCTTTCTTAAGAGTGTATTT 765

Qy     303 ATTACAA 309
           ||  |||
Db     766 ATGCCAA 772


RESULT 12
ABL61761
ID   ABL61761 standard; DNA; 5544 BP.
XX
AC   ABL61761;
XX
DT   15-MAY-2002  (first entry)
XX
DE   Colon adenocarcinoma related gene sequence SEQ ID NO:98.
XX
KW   Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW   stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW   cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW   gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200194629-A2.
XX
PD   13-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US10838.
XX
PR   05-JUN-2000; 2000US-209473P.
```

PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
PS Claim 1; SEQ ID 98; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX

SQ Sequence 5544 BP; 1504 A; 1281 C; 1157 G; 1602 T; 0 other;

Query Match 10.3%; Score 36.6; DB 24; Length 5544;
Best Local Similarity 49.7%; Pred. No. 0.67;
Matches 93; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

```
Qy    123 TGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGC 182
          |||   |   |  ||||||    | |    ||  |  |  | || | |  | | ||  ||
Db   2106 TGACCCCCTGGCACCAACACCCGCCAGTACTGTGACTTCCAAAGCCAGAGCCACATGTGC 2165

Qy    183 TGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAG 242
          | ||   | | | |  || ||    ||    |   || ||||  ||||||||     |
Db   2166 TCATCAAACTTGCATTAAGCAGTTGGCGGGAGATGGCTGTGGAGCTGGGGGTTTAAGTGA 2225

Qy    243 TAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCACT 302
          | ||    ||  ||  |  ||| |   | ||    |  | ||| | |||| |||||   |
Db   2226 TGGTTCTCTTTTGCTCCCTCTTTTGAGGGTAAAGCTACTGTCTTTCTTAAGAGTGTATTT 2285

Qy    303 ATTACAA 309
          ||  |||
Db   2286 ATGCCAA 2292
```

RESULT 13
AAS87055
ID AAS87055 standard; cDNA; 5549 BP.
XX
AC AAS87055;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #22859.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG22868.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 22859; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.

```
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 5549 BP; 1510 A; 1285 C; 1154 G; 1600 T; 0 other;

  Query Match            10.3%;  Score 36.6;  DB 23;  Length 5549;
  Best Local Similarity  49.7%;  Pred. No. 0.67;
  Matches   93;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy     123 TGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGC 182
           |||   |   |  ||||||    | |    ||  |  |  | || | |  | | ||  ||
Db    2106 TGACCCCCTGGCACCAACACCCGCCAGTACTGTGACTTCCAAAGCCAGAGCCACATGTGC 2165

Qy     183 TGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAG 242
           | ||   | | | |  || ||    ||    |  || ||||  ||||||||     |
Db    2166 TCATCAAACTTGCATTAAGCAGTTGGCGGGAGATGGCTGTGGAGCTGGGGGTTTAAGTGA 2225

Qy     243 TAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCACT 302
           | ||    ||  ||  |  ||| |   | ||    |  | ||| | |||| |||||   |
Db    2226 TGGTTCTCTTTTGCTCCCTCTTTTGAGGGTAAAGCTACTGTCTTTCTTAAGAGTGTATTT 2285

Qy     303 ATTACAA 309
           ||  |||
Db    2286 ATGCCAA 2292


RESULT 14
ABI99688/c
ID   ABI99688 standard; cDNA; 2215 BP.
XX
AC   ABI99688;
XX
DT   07-MAR-2002  (first entry)
XX
DE   Mouse ischaemic condition related cDNA sequence SEQ ID NO:726.
XX
KW   Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW   vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
XX
OS   Mus musculus.
XX
PN   WO200188188-A2.
XX
PD   22-NOV-2001.
XX
PF   18-MAY-2001; 2001WO-JP04192.
XX
PR   18-MAY-2000; 2000JP-0145977.
XX
PA   (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI   Ishikawa K,  Asai S,  Takahashi Y,  Nagata T,  Ishii Y;
XX
DR   WPI; 2002-034733/04.
DR   P-PSDB; ABB57259.
XX
PT   Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT   expression levels of particular genes defined in the specification or
PT   by determining the expression profile of a gene group comprising these
PT   genes -
XX
PS   Claim 2; Page 1794-1796; 2690pp; English.
XX
CC   The present invention describes a method for examining ischaemic
CC   conditions, comprising measuring the expression levels of particular
CC   genes (I) in a test sample or determining the expression profile of a
CC   gene group in the sample comprising genes selected from (I). The method
CC   is useful for examining the ischaemic condition (e.g. compressive
CC   ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC   expression levels of particular genes (ABI99202 to ABI99912, encoding
CC   the protein sequences in ABB57020 to ABB57374) or by determining the
CC   expression profile of a gene group comprising these genes. The
CC   expression levels or expression profiles produced by these genes are
CC   used as an indicator when screening for ischaemic condition-improving
CC   drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914
CC   represent PCR primers for a mouse ischaemic condition related sequence,
CC   which are used in the exemplification of the present invention.
XX
SQ   Sequence 2215 BP; 640 A; 591 C; 457 G; 527 T; 0 other;

  Query Match            10.2%;  Score 36.4;  DB 24;  Length 2215;
  Best Local Similarity  46.7%;  Pred. No. 0.53;
  Matches  115;  Conservative    0;  Mismatches  131;  Indels    0;  Gaps    0;

Qy      41 CTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAG 100
           || ||| || || || | ||| ||    |     ||   |   ||          || |
Db    1630 CTGTTGCTGCTGCTGGTGGTGGTGGTCATGGAACTGCTGTTGCTGCTGGGGGTGGTCATG 1571

Qy     101 CAGCTGCGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATAT 160
           | ||||  |   | ||||| ||| | | |||    |  |   | ||   ||||||
Db    1570 GAACTGCTGCTTCTGCTGGGAGTGGTCATGGAACTGCTGCTGCTGCTGCTGGTGGTGGTC 1511

Qy     161 ATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCG 220
           || ||   | | |     | |||| |  |  |  |   | |  |  ||||   ||  | |
Db    1510 ATGGAACTGCTGCTGCTGCTGCTGGTGGTGGTCATGGAACTGCTGCTGCTGTTGGTGGTG 1451

Qy     221 GTGGGTCTGGGGGTAGCGTTAGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTG 280
           ||||   |   || |  | |  |  | |||    ||    | ||    |||| ||
Db    1450 GTGGTGGTCATGGAACTGCTGCTTCTGCTGGTGGTGGTCATGGAACTGCTGCTGCTACTG 1391

Qy     281 TTTCTG 286
           | |||
Db    1390 CTGCTG 1385


RESULT 15
ABK40083
ID   ABK40083 standard; DNA; 1668 BP.
XX
AC   ABK40083;
XX
DT   21-MAY-2002  (first entry)
XX
DE   Human chemically pretreated gene sequence #83 strand 1.
XX
KW   Human; ds; bisulphite treatment; CpG; DNA methylation; cancer; tumour;
KW   cytostatic; ALDH6; CYP11A; CYP11B1; CYP3A3; DPYD; EPHX2; OCLN; TXNRD1;
KW   UGT8; MRP; pharmacogenomics; SNP; single nucleotide polymorphism.
XX
OS   Homo sapiens.
XX
PN   WO200202806-A2.
XX
PD   10-JAN-2002.
XX
PF   29-JUN-2001; 2001WO-EP07470.
XX
PR   30-JUN-2000; 2000DE-1032529.
PR   01-SEP-2000; 2000DE-1043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-154757/20.
XX
PT   New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,
PT   useful for detecting cytosine methylation state of genes associated
PT   with pharmacogenomics and for therapy of diseases e.g. cancer  -
```

XX
PS Claim 1; SEQ ID No 165; 24pp; English.
XX
CC The invention relates to a nucleic acid comprising a sequence at
CC least 18 bases in length of a segment of the chemically pretreated DNA
CC of genes associated with pharmacogenomics according to one of the
CC sequences of the genes ALDH6 (NM_000693), CYP11A (NM_000781), CYP11B1
CC (NM_000497), CYP3A3 (NM_000776 and NM_017460), DPYD (NM_000110), EPHX2
CC (NM_001979), OCLN (NM_002538), TXNRD1 (NM_003330), UGT8 (NM_003360),
CC MRP (NM_004996, NM_019900, NM_019901, NM_019902, NM_019862, NM_019898,
CC NM_019899) and their complementary sequences, or a sequence (S1) chosen
CC from 87 sequences and their complements. The chemical pretreatment
CC is bisulphite treatment to convert cytosines (but not methyl-cytosines)
CC into uracils. Also included are an oligomer (II) in particular an
CC oligonucleotide or a peptide nucleic acid (PNA)-oligomer, comprising in
CC each case at least one base sequence having a length of 9 nucleotides
CC which hybridises to or is identical to a chemically pretreated DNA of
CC genes associated with pharmacogenomics and their complements, arranged in
CC an array for analysing diseases associated with the methylation state
CC (CpG) and/or detecting SNPs (single nucleotide polymorphisms)
CC of the 87 sequences. The oligomers may also be used as PCR primers.
CC The set of 87 nucleic acids and their complements is useful for diagnosis
CC and therapy of solid tumours and cancer. The present sequence
CC represents one the 87 DNA sequences or its complement.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1668 BP; 480 A; 15 C; 359 G; 814 T; 0 other;

Query Match 10.0%; Score 35.6; DB 24; Length 1668;
Best Local Similarity 51.2%; Pred. No. 0.84;
Matches 83; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

```
Qy    141 ATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAA 200
          ||||| ||  |  ||||||| |||  | | |    |   |  |  ||||| | |||  ||
Db    387 ATTTTTTATTGTGGGTATATTTAGGTAAGTTTTTTGTGTAAGTTTTTTTATTTGTATAAA 446

Qy    201 TCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAGTAGTCGAGTTGGGCAGCA 260
            ||   |  || |  |  || |  | || || | | || |  | | |    ||   |
Db    447 ATATTTAGTGTAAGTATTTGGGGTTTTTGTGGATTGGGTAATGATTTGGAAGGGTTGGAG 506

Qy    261 ACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCACT 302
            || ||| |     | || ||  ||| |   || ||| | |
Db    507 GTTTTTTGGGGATTTTTTTTTTATTGTTTGTTTAATGTAAGT 548
```

Search completed: March 12, 2003, 08:23:48
Job time : 566 secs

OM nucleic - nucleic search, using sw model

Run on: March 11, 2003, 10:26:12 ; Search time 53.5 Seconds
(without alignments)
2046.423 Million cell updates/sec

Title: US-09-936-011-3
Perfect score: 357
Sequence: 1 atgtctagggaaataaccgc.....gtaatggtgcttgtggttaa 357

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	30.4	8.5	9595	3	US-09-014-416-4	Sequence 4, Appli
c	2	29.6	8.3	2087	4	US-09-097-199-83	Sequence 83, Appl
c	3	29.6	8.3	2505	4	US-09-097-199-85	Sequence 85, Appl
	4	29.6	8.3	4285	4	US-09-410-464-1	Sequence 1, Appli
	5	29	8.1	1370	4	US-09-026-408-12	Sequence 12, Appl
	6	29	8.1	1371	4	US-09-026-408-1	Sequence 1, Appli
c	7	28.8	8.1	5163	3	US-08-700-651-1	Sequence 1, Appli
c	8	28.8	8.1	5163	3	US-08-928-361B-4	Sequence 4, Appli
c	9	28.8	8.1	5318	3	US-08-700-651-2	Sequence 2, Appli
c	10	28.8	8.1	5318	3	US-08-928-361B-3	Sequence 3, Appli
c	11	28.8	8.1	5883	1	US-08-404-445-1	Sequence 1, Appli
	12	28.8	8.1	6060	5	PCT-US96-09430-7	Sequence 7, Appli
c	13	28.8	8.1	8442	4	US-09-272-032-6	Sequence 6, Appli
	14	28.6	8.0	1506	2	US-08-663-566A-8	Sequence 8, Appli
	15	28.6	8.0	1506	2	US-08-023-610-8	Sequence 8, Appli
	16	28.6	8.0	1506	2	US-08-288-065A-8	Sequence 8, Appli
	17	28.6	8.0	1506	2	US-08-362-240A-8	Sequence 8, Appli
	18	28.6	8.0	1506	4	US-08-804-372A-6	Sequence 6, Appli
	19	28.6	8.0	1506	5	PCT-US95-10245-8	Sequence 8, Appli
c	20	28.4	8.0	660	4	US-09-549-804C-1	Sequence 1, Appli
	21	28.4	8.0	3454	4	US-08-961-527-207	Sequence 207, App
c	22	28.4	8.0	5173	4	US-08-801-308-2	Sequence 2, Appli
c	23	28.4	8.0	16569	4	US-09-097-889-2	Sequence 2, Appli
c	24	28.4	8.0	16569	4	US-09-377-856-1	Sequence 1, Appli
c	25	28.4	8.0	16569	4	US-09-302-681-2	Sequence 2, Appli
	26	28.4	8.0	18596	4	US-09-318-448-11	Sequence 11, Appl
	27	28.2	7.9	1170	5	PCT-US96-05320A-900	Sequence 900, App
c	28	28.2	7.9	1316	3	US-08-460-576-1	Sequence 1, Appli
c	29	28.2	7.9	5511	3	US-08-928-361B-2	Sequence 2, Appli
c	30	28.2	7.9	7334	3	US-08-928-361B-1	Sequence 1, Appli
c	31	28	7.8	549	1	US-08-196-945-1	Sequence 1, Appli
	32	28	7.8	7766	4	US-09-125-619-3	Sequence 3, Appli
c	33	27.8	7.8	3134	2	US-08-533-669A-1	Sequence 1, Appli
c	34	27.8	7.8	3134	4	US-09-183-861-1	Sequence 1, Appli
c	35	27.8	7.8	3134	4	US-09-022-765-1	Sequence 1, Appli
	36	27.6	7.7	1251	2	US-08-655-704B-16	Sequence 16, Appl
	37	27.6	7.7	1251	3	US-09-107-755-16	Sequence 16, Appl
	38	27.6	7.7	2265	3	US-09-369-618-3	Sequence 3, Appli
	39	27.6	7.7	2265	3	US-09-369-617-3	Sequence 3, Appli
	40	27.4	7.7	377	2	US-08-332-766A-1	Sequence 1, Appli
c	41	27.4	7.7	1022	1	US-08-198-446B-3	Sequence 3, Appli
c	42	27.4	7.7	1022	2	US-08-870-693-3	Sequence 3, Appli
c	43	27.4	7.7	3054	4	US-09-484-970B-138	Sequence 138, App
	44	27.4	7.7	4310	3	US-09-008-172-1	Sequence 1, Appli
	45	27.4	7.7	4310	4	US-09-210-361-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-014-416-4
; Sequence 4, Application US/09014416
; Patent No. 6153421
; GENERAL INFORMATION:
; APPLICANT: Yanagi, Masayuki
; APPLICANT: Bukh, Jens
; APPLICANT: Emerson, Susanne U.
; APPLICANT: Purcell, Robert H.
; TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 20264276
; CURRENT APPLICATION NUMBER: US/09/014,416
; CURRENT FILING DATE: 1998-01-27
; EARLIER APPLICATION NUMBER: US 60/053,062
; EARLIER FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 9595
; TYPE: DNA
; ORGANISM: Hepatitis C virus
US-09-014-416-4

Query Match 8.5%; Score 30.4; DB 3; Length 9595;
Best Local Similarity 61.2%; Pred. No. 6.1;
Matches 49; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

```
Qy       1 ATGTCTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGT 60
           |||    ||||  ||||| ||| |  |  ||||    | ||| || || ||  | ||| |
Db    9371 ATGAACGGGGAGCTAACCACTCCAGGCCTTAAGCCATTTCCTGTTTTTTTTTTTTTTTT 9430

Qy      61 GTTGTGGCTTTCTTTGTATT 80
             || |   ||||||| | ||
Db    9431 TTTTTTTTTTCTTTTTTT 9450
```

RESULT 2
US-09-097-199-83/c
; Sequence 83, Application US/09097199
; Patent No. 6218529
; GENERAL INFORMATION:
; APPLICANT: An, Gang
; APPLICANT: O'Hara, S. Mark
; APPLICANT: Ralph, David
; APPLICANT: Veltri, Robert
; TITLE OF INVENTION: BIOMARKERS AND TARGETS FOR DIAGNOSIS,
; TITLE OF INVENTION: PROGNOSIS AND MANAGEMENT OF PROSTATE DISEASE
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:

```
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/097,199
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/692,787
;      FILING DATE:  31-JUL-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Nakashima, Richard A.
;      REGISTRATION NUMBER:  P-42,023
;      REFERENCE/DOCKET NUMBER:  UROC:018
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  83:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2087 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  99..503
US-09-097-199-83

  Query Match             8.3%;  Score 29.6;  DB 4;  Length 2087;
  Best Local Similarity   46.6%;  Pred. No. 5.4;
  Matches   95;  Conservative    0;  Mismatches  109;  Indels    0;  Gaps    0;

Qy     103 GCTGCGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATAT 162
           | || || |  | | |  || |||| |      |   |  ||     ||| || |   | |
Db     346 GGTGTGAGAATTGCATGGTGTGACTGTTCTATGGAATAATCAATGAACTCTGTTGCCTTT 287

Qy     163 AGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGT 222
             ||     |   |  |   ||  ||||   | || |    | |     |      | ||
Db     286 TGACTTCTTTTTACTTTTTCTTTTTTTTCGAAATACCTTGTAGCAGAATGGTAAGAGCACA 227

Qy     223 GGGTCTGGGGGTAGCGTTAGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTT 282
           |  ||||| | |||  |   | || |  ||||| ||  ||||| | |||     | | ||
Db     226 GTTTCTGGAGCTAGAATGCCTGGTAGCCTTGGGAAGTTACTTAATCTGTGCCTCAATTTT 167

Qy     283 TCTGTGTTAATAGTGTCACTATTA 306
               | |  | | ||| | || ||
Db     166 CTGATCTGTAAAATGTGAATAATA 143


RESULT 3
US-09-097-199-85/c
; Sequence 85, Application US/09097199
; Patent No. 6218529
;  GENERAL INFORMATION:
;    APPLICANT:  An, Gang
;    APPLICANT:  O'Hara, S. Mark
;    APPLICANT:  Ralph, David
;    APPLICANT:  Veltri, Robert
;    TITLE OF INVENTION:  BIOMARKERS AND TARGETS FOR DIAGNOSIS,
;    TITLE OF INVENTION:  PROGNOSIS AND MANAGEMENT OF PROSTATE DISEASE
;    NUMBER OF SEQUENCES:  87
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/097,199
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/692,787
;      FILING DATE:  31-JUL-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Nakashima, Richard A.
;      REGISTRATION NUMBER:  P-42,023
;      REFERENCE/DOCKET NUMBER:  UROC:018
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  85:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2505 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  99..503
US-09-097-199-85

  Query Match             8.3%;  Score 29.6;  DB 4;  Length 2505;
  Best Local Similarity   46.6%;  Pred. No. 5.9;
  Matches   95;  Conservative    0;  Mismatches  109;  Indels    0;  Gaps    0;

Qy     103 GCTGCGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATAT 162
           | || || |  | | |  || |||| |      |   |  ||     ||| || |   | |
Db     346 GGTGTGAGAATTGCATGGTGTGACTGTTCTATGGAATAATCAATGAACTCTGTTGCCTTT 287

Qy     163 AGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGT 222
             ||     |   |  |   ||  ||||   | || |    | |     |      | ||
Db     286 TGACTTCTTTTTACTTTTTCTTTTTTTTCGAAATACCTTGTAGCAGAATGGTAAGAGCACA 227

Qy     223 GGGTCTGGGGGTAGCGTTAGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTT 282
           |  ||||| | |||  |   | || |  ||||| ||  ||||| | |||     | | ||
Db     226 GTTTCTGGAGCTAGAATGCCTGGTAGCCTTGGGAAGTTACTTAATCTGTGCCTCAATTTT 167

Qy     283 TCTGTGTTAATAGTGTCACTATTA 306
               | |  | | ||| | || ||
Db     166 CTGATCTGTAAAATGTGAATAATA 143


RESULT 4
US-09-410-464-1
; Sequence 1, Application US/09410464
; Patent No. 6395892
; GENERAL INFORMATION:
;  APPLICANT: Strauss et al.
;  TITLE OF INVENTION: Floral homeotic genes for manipulation of flowering in
;  TITLE OF INVENTION: poplar and other plant species.
;  FILE REFERENCE: 53375
;  CURRENT APPLICATION NUMBER: US/09/410,464
;  CURRENT FILING DATE: 1999-10-01
;  EARLIER APPLICATION NUMBER: 09/287,700
;  EARLIER FILING DATE: 1999-04-06
;  EARLIER APPLICATION NUMBER: 60/080,851
;  EARLIER FILING DATE: 1998-04-06
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 4285
```

```
;   TYPE: DNA
;   ORGANISM: Populus balsamifera subsp. trichocarpa
US-09-410-464-1

  Query Match              8.3%;  Score 29.6;  DB 4;  Length 4285;
  Best Local Similarity   54.6%;  Pred. No. 7.6;
  Matches   59;  Conservative    0;  Mismatches   49;  Indels    0;  Gaps    0;

Qy     175 AGATCAGCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGT 234
           | ||||| | | || || || || || | |||    || |  || |||     |      |
Db    2540 AAATCAGGTAACTTCAAAAGAAATAACCTTCGCATATATGCATGTGGTTATGGTTTTTAT 2599

Qy     235 AGCGTTAGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTT 282
           |   ||   ||  | ||| | ||| ||| |||  | ||  | || ||
Db    2600 GGGAATATCTGTAAATTTGTGGAGCTACTAATTAAGGTATTTGTTTTT 2647


RESULT 5
US-09-026-408-12
; Sequence 12, Application US/09026408
; Patent No. 6303338
;  GENERAL INFORMATION:
;    APPLICANT:  Ni et al.
;    TITLE OF INVENTION:  PANCREAS-DERIVED PLASMINOGEN ACTIVATOR
;    TITLE OF INVENTION:  INHIBITOR
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;      STREET:  1100 NEW YORK AVENUE, SUITE 600
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/026,408
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/934,011
;      FILING DATE:  15-AUG-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/024,056
;      FILING DATE:  16-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  STEFFE, ERIC K.
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1488.0300002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-2600
;      TELEFAX:  202-371-2540
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1370 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  67..1281
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  67..120
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  121..1281
US-09-026-408-12

  Query Match              8.1%;  Score 29;  DB 4;  Length 1370;
  Best Local Similarity   54.1%;  Pred. No. 6.9;
  Matches   59;  Conservative    0;  Mismatches   50;  Indels    0;  Gaps    0;

Qy     241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
           || |||| |  ||| ||  | ||| |||||    ||  | || |||| ||      ||||
Db      57 AGAAGTCAAAATGGACACAATCTTCTTGTGGAGTCTTCTATTGCTGTTTTTTGGAAGTCA 116

Qy     301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTT 349
           |  |     | || ||   |  | | |  ||||| | |||  |||
Db     117 AGCCTCAAGATGCTCAGCTCAAAAAAATACCGAATTTGCAGTGGATCTT 165


RESULT 6
US-09-026-408-1
; Sequence 1, Application US/09026408
; Patent No. 6303338
;  GENERAL INFORMATION:
;    APPLICANT:  Ni et al.
;    TITLE OF INVENTION:  PANCREAS-DERIVED PLASMINOGEN ACTIVATOR
;    TITLE OF INVENTION:  INHIBITOR
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;      STREET:  1100 NEW YORK AVENUE, SUITE 600
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/026,408
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/934,011
;      FILING DATE:  15-AUG-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/024,056
;      FILING DATE:  16-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  STEFFE, ERIC K.
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1488.0300002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-2600
;      TELEFAX:  202-371-2540
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1371 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  both
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  67..1242
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  109..1242
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  67..108
US-09-026-408-1

  Query Match              8.1%;  Score 29;  DB 4;  Length 1371;
  Best Local Similarity   54.1%;  Pred. No. 6.9;
  Matches   59;  Conservative    0;  Mismatches   50;  Indels    0;  Gaps    0;
```

```
Qy    241 AGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCA 300
          || |||| |  ||| ||  | ||| |||||    ||  | || |||| ||      ||||
Db     57 AGAAGTCAAAATGGACACAATCTTCTTGTGGAGTCTTCTATTGCTGTTTTTTGGAAGTCA 116

Qy    301 CTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTT 349
            |  |     | || ||   |  | | |  ||||| | |||  |||
Db    117 AGCCTCAAGATGCTCAGCTCAAAAAAATACCGAATTTGCAGTGGATCTT 165


RESULT 7
US-08-700-651-1/c
; Sequence 1, Application US/08700651B
; Patent No. 6015882
; GENERAL INFORMATION:
;  APPLICANT: PETERSEN, CAROLYN
;  APPLICANT: LEECH, JAMES
;  APPLICANT: NELSON, RICHARD, C.
;  APPLICANT: GUT, JIRI
;  TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS
;  TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
;  TITLE OF INVENTION: INFECTIONS
;  FILE REFERENCE: 480.19-4(HV)
;  CURRENT APPLICATION NUMBER: US/08/700,651B
;  CURRENT FILING DATE: 1997-08-14
;  EARLIER APPLICATION NUMBER: 08/415,751
;  EARLIER FILING DATE: 1995-04-03
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;    LENGTH: 5163
;    TYPE: DNA
;    ORGANISM: Cryptosporidium parvum
US-08-700-651-1

  Query Match             8.1%;  Score 28.8;  DB 3;  Length 5163;
  Best Local Similarity   49.6%;  Pred. No. 15;
  Matches  130;  Conservative    0;  Mismatches  127;  Indels    5;  Gaps    2;

Qy     36 TGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCA 95
          |||   | |||||||||  |||  |||||| | |||||| ||| |  |||   |  |
Db    952 TGTTGTTGTTGTTGTTGTAGTTGTTGTTGTTGGTTTCTTGGTAGTTGTGGTTGTGGTTGT 893

Qy     96 GCAAGCAGCTGCGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACGGTGG 155
          |   |  |  |    |  | || || ||     |  ||   |     || |   ||| |
Db    892 GGTTGTTGTGGTTGTAGTTGTTGTGGTTGTAGTTGTGGTTGTA----GTTGTTGTGGTTG 837

Qy    156 TATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCGTGCTTACGG 215
          |   | | | |   |||   | |  | |  |  |  | ||   | |  | ||| ||   |
Db    836 TTGTTGTTGTGGTTGTAGTTGTTGTGGTTGTAGTTGTGGTTGTAGTTGTTGTGGTTGTAG 777

Qy    216 GTGCGGTGGGTCTGGGGGTAGCGTTAGTAGTCG-AGTTGGGCAGCAACTTATTGTGTTAG 274
            |  || |   | |  ||||   ||||||| | ||| |          || ||||  | |
Db    776 TAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTCGTTGTTGTTGTTGTTGTTG 717

Qy    275 CTATTGTTTCTGTGTTAATAGT 296
           | |||||  |||  |  | ||
Db    716 TTGTTGTTGTTGTTGTTGTTGT 695


RESULT 8
US-08-928-361B-4/c
; Sequence 4, Application US/08928361B
; Patent No. 6071518
;  GENERAL INFORMATION:
;    APPLICANT:  Petersen, Carolyn
;    TITLE OF INVENTION:  PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
;    TITLE OF INVENTION:  THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
;    TITLE OF INVENTION:  FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
;    TITLE OF INVENTION:  SPECIES INFECTIONS
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PETERS, VERNY, JONES & BIKSA
;      STREET:  385 Sherman Avenue, Suite 6
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306-1840
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/928,361B
;      FILING DATE:  12-SEP-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/026,062
;      FILING DATE:  13-SEP-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verny, Hana
;      REGISTRATION NUMBER:  30,518
;      REFERENCE/DOCKET NUMBER:  480.76-1(HV)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-324-1677
;      TELEFAX:  650-324-1678
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5163 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-928-361B-4

  Query Match             8.1%;  Score 28.8;  DB 3;  Length 5163;
  Best Local Similarity   49.6%;  Pred. No. 15;
  Matches  130;  Conservative    0;  Mismatches  127;  Indels    5;  Gaps    2;

Qy     36 TGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCA 95
          |||   | |||||||||  |||  |||||| | |||||| ||| |  |||   |  |
Db    951 TGTTGTTGTTGTTGTTGTAGTTGTTGTTGTTGGTTTCTTGGTAGTTGTGGTTGTGGTTGT 892

Qy     96 GCAAGCAGCTGCGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACGGTGG 155
          |   |  |  |    |  | || || ||     |  ||   |     || |   ||| |
Db    891 GGTTGTTGTGGTTGTAGTTGTTGTGGTTGTAGTTGTGGTTGTA----GTTGTTGTGGTTG 836

Qy    156 TATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCGTGCTTACGG 215
          |   | | | |   |||   | |  | |  |  |  | ||   | |  | ||| ||   |
Db    835 TTGTTGTTGTGGTTGTAGTTGTTGTGGTTGTAGTTGTGGTTGTAGTTGTTGTGGTTGTAG 776

Qy    216 GTGCGGTGGGTCTGGGGGTAGCGTTAGTAGTCG-AGTTGGGCAGCAACTTATTGTGTTAG 274
            |  || |   | |  ||||   ||||||| | ||| |          || ||||  | |
Db    775 TAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTCGTTGTTGTTGTTGTTGTTG 716

Qy    275 CTATTGTTTCTGTGTTAATAGT 296
           | |||||  |||  |  | ||
Db    715 TTGTTGTTGTTGTTGTTGTTGT 694


RESULT 9
US-08-700-651-2/c
; Sequence 2, Application US/08700651B
; Patent No. 6015882
; GENERAL INFORMATION:
;  APPLICANT: PETERSEN, CAROLYN
;  APPLICANT: LEECH, JAMES
;  APPLICANT: NELSON, RICHARD, C.
;  APPLICANT: GUT, JIRI
;  TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS
;  TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
;  TITLE OF INVENTION: INFECTIONS
;  FILE REFERENCE: 480.19-4(HV)
;  CURRENT APPLICATION NUMBER: US/08/700,651B
;  CURRENT FILING DATE: 1997-08-14
;  EARLIER APPLICATION NUMBER: 08/415,751
```

```
;  EARLIER FILING DATE: 1995-04-03
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;    LENGTH: 5318
;    TYPE: DNA
;    ORGANISM: Cryptosporidium parvum
US-08-700-651-2

  Query Match             8.1%;  Score 28.8;  DB 3;  Length 5318;
  Best Local Similarity   49.6%;  Pred. No. 15;
  Matches  130;  Conservative    0;  Mismatches  127;  Indels    5;  Gaps    2;

Qy      36 TGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCA 95
           |||   | |||||||||  |||  |||||| | |||||| ||| |  |||   |  |
Db     952 TGTTGTTGTTGTTGTTGTAGTTGTTGTTGTTGGTTTCTTGGTAGTTGTGGTTGTGGTTGT 893

Qy      96 GCAAGCAGCTGCGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACGGTGG 155
           |   |  |  |     |  |  || || ||     |  ||    |     || |   ||| |
Db     892 GGTTGTTGTGGTTGTAGTTGTTGTGGTTGTAGTTGTGGTTGTA----GTTGTTGTGGTTG 837

Qy     156 TATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCGTGCTTACGG 215
           |   | | | |   |||   | |  | |  |  | ||   | |  | ||| ||   |
Db     836 TTGTTGTTGTGGTTGTAGTTGTTGTGGTTGTAGTTGTGGTTGTAGTTGTTGTGGTTGTAG 777

Qy     216 GTGCGGTGGGTCTGGGGGTAGCGTTAGTAGTCG-AGTTGGGCAGCAACTTATTGTGTTAG 274
             |  || |   | |  ||||   |||||||  | ||| |          || ||||  | |
Db     776 TAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTCGTTGTTGTTGTTGTTGTTG 717

Qy     275 CTATTGTTTCTGTGTTAATAGT 296
           | |||||  |||  |  | ||
Db     716 TTGTTGTTGTTGTTGTTGTTGT 695


RESULT 10
US-08-928-361B-3/c
; Sequence 3, Application US/08928361B
; Patent No. 6071518
;  GENERAL INFORMATION:
;    APPLICANT:  Petersen, Carolyn
;    TITLE OF INVENTION:  PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
;    TITLE OF INVENTION:  THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
;    TITLE OF INVENTION:  FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
;    TITLE OF INVENTION:  SPECIES INFECTIONS
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PETERS, VERNY, JONES & BIKSA
;      STREET:  385 Sherman Avenue, Suite 6
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306-1840
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/928,361B
;      FILING DATE:  12-SEP-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/026,062
;      FILING DATE:  13-SEP-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Verny, Hana
;      REGISTRATION NUMBER:  30,518
;      REFERENCE/DOCKET NUMBER:  480.76-1(HV)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-324-1677
;      TELEFAX:  650-324-1678
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5318 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-928-361B-3

  Query Match             8.1%;  Score 28.8;  DB 3;  Length 5318;
  Best Local Similarity   49.6%;  Pred. No. 15;
  Matches  130;  Conservative    0;  Mismatches  127;  Indels    5;  Gaps    2;

Qy      36 TGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCA 95
           |||   | |||||||||  |||  |||||| | |||||| ||| |  |||   |  |
Db     951 TGTTGTTGTTGTTGTTGTAGTTGTTGTTGTTGGTTTCTTGGTAGTTGTGGTTGTGGTTGT 892

Qy      96 GCAAGCAGCTGCGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACGGTGG 155
           |   |  |  |     |  |  || || ||     |  ||    |     || |   ||| |
Db     891 GGTTGTTGTGGTTGTAGTTGTTGTGGTTGTAGTTGTGGTTGTA----GTTGTTGTGGTTG 836

Qy     156 TATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCGTGCTTACGG 215
           |   | | | |   |||   | |  | |  |  | ||   | |  | ||| ||   |
Db     835 TTGTTGTTGTGGTTGTAGTTGTTGTGGTTGTAGTTGTGGTTGTAGTTGTTGTGGTTGTAG 776

Qy     216 GTGCGGTGGGTCTGGGGGTAGCGTTAGTAGTCG-AGTTGGGCAGCAACTTATTGTGTTAG 274
             |  || |   | |  ||||   |||||||  | ||| |          || ||||  | |
Db     775 TAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTCGTTGTTGTTGTTGTTGTTG 716

Qy     275 CTATTGTTTCTGTGTTAATAGT 296
           | |||||  |||  |  | ||
Db     715 TTGTTGTTGTTGTTGTTGTTGT 694


RESULT 11
US-08-404-445-1/c
; Sequence 1, Application US/08404445
; Patent No. 5753220
;  GENERAL INFORMATION:
;    APPLICANT:  SUZUKI, TAKEO
;    APPLICANT:  USAMI, AKIHIRO
;    APPLICANT:  ODA, KOUHEI
;    APPLICANT:  MORI, HAJIME
;    APPLICANT:  KANAYA, TOSHIMICHI
;    TITLE OF INVENTION:  CYSTEINE PROTEASE GENE DEFECTIVE
;    TITLE OF INVENTION:  BACULOVIRUS, PROCESS FOR ITS PRODUCTION, AND PROCESS FOR
;    TITLE OF INVENTION:  THE PRODUCTION OF ECONOMIC PROTEIN BY USING THE SAME
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
;      STREET:  1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
;      CITY:  ARLINGTON
;      STATE:  VIRGINIA
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/404,445
;      FILING DATE:  15-MAR-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 68961/1994
;      FILING DATE:  15-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 23296/1995
;      FILING DATE:  19-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OBLON, NORMAN F
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  4508-002-0
;    TELECOMMUNICATION INFORMATION:
```

```
;       TELEPHONE:  (703) 413-3000
;       TELEFAX:  (703) 413-2220
;       TELEX:  248855 OPAT UR
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  5883 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     ORIGINAL SOURCE:
;       ORGANISM:  SILKWORM NUCLEAR POLYHEDROSIS VIRUS
;       STRAIN:  BmNPVCP
;     FEATURE:
;       NAME/KEY:  mat_peptide
;       LOCATION:  1032..4642
US-08-404-445-1

  Query Match             8.1%;  Score 28.8;  DB 1;  Length 5883;
  Best Local Similarity   65.6%;  Pred. No. 16;
  Matches   42;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy     205 CGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAGTAGTCGAGTTGGGCAGCAACTT 264
           | |||| |||| | | || | |||| ||| || | ||||  ||||    |||| | ||
Db    3768 CATGCTGACGGTTACCGTTGTTCTGCGGGAAGGGATAGTTCTCGAACATGGCATCGACGG 3709

Qy     265 ATTG 268
           |||
Db    3708 CTTG 3705


RESULT 12
PCT-US96-09430-7
; Sequence 7, Application PC/TUS9609430
;  GENERAL INFORMATION:
;     APPLICANT:  Glazer, Peter M.
;     TITLE OF INVENTION:  TREATMENT OF HEMOGLOBINOPATHIES
;     NUMBER OF SEQUENCES:  23
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  OncorPharm, Inc.
;       STREET:  200 Perry Parkway
;       CITY:  Gaithersburg
;       STATE:  Maryland
;       COUNTRY:  US
;       ZIP:  20877
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US96/09430
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/473,845
;       FILING DATE:  07-JUN-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Karta, Glenn E.
;       REGISTRATION NUMBER:  30,649
;       REFERENCE/DOCKET NUMBER:  PA-0040
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  301-527-2058
;       TELEFAX:  301-208-6997
;   INFORMATION FOR SEQ ID NO:  7:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  6060 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     ORIGINAL SOURCE:
;       ORGANISM:  Homo sapiens
;     POSITION IN GENOME:
;       CHROMOSOME/SEGMENT:  11
PCT-US96-09430-7

  Query Match             8.1%;  Score 28.8;  DB 5;  Length 6060;
  Best Local Similarity   53.6%;  Pred. No. 16;
  Matches   60;  Conservative    0;  Mismatches   52;  Indels    0;  Gaps    0;

Qy     245 GTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCACTAT 304
           || |||   || || ||      | | |  |||||  ||  |   |||||| |  |  ||
Db     553 GTTGAGACAGGTAGAAAAAGTGAGAGATCTCTATTTATTTAGCAATAATAGAGAAAGCAT 612

Qy     305 TACAACGATTAAGGTCTCCACCAGAACACATTTGTAATGGTGCTTGTGGTTA 356
           |  |  || ||| |         | || | ||||||||   | ||| || | |
Db     613 TTAAGAGAATAAAGCAATGGAAATAAGAAATTTGTAAATTTCCTTCTGATAA 664


RESULT 13
US-09-272-032-6/c
; Sequence 6, Application US/09272032A
; Patent No. 6296852
; GENERAL INFORMATION:
;  APPLICANT: Johnson, Michael A.
;  APPLICANT: Prideaux, Christopher T.
;  APPLICANT: McCoy, Richard J.
;  TITLE OF INVENTION: Recombinant Avian Adenovirus Vector
;  FILE REFERENCE: 48-95a
;  CURRENT APPLICATION NUMBER: US/09/272,032A
;  CURRENT FILING DATE: 1999-03-18
;  EARLIER APPLICATION NUMBER: US 08/448,617
;  EARLIER FILING DATE: 1995-09-08
;  EARLIER APPLICATION NUMBER: PCT/AU94/000189
;  EARLIER FILING DATE: 1994-04-14
;  EARLIER APPLICATION NUMBER: AU PL 8297
;  EARLIER FILING DATE: 1993-04-14
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
;   LENGTH: 8442
;   TYPE: DNA
;   ORGANISM: fowl adenovirus
US-09-272-032-6

  Query Match             8.1%;  Score 28.8;  DB 4;  Length 8442;
  Best Local Similarity   53.6%;  Pred. No. 19;
  Matches   60;  Conservative    0;  Mismatches   52;  Indels    0;  Gaps    0;

Qy     144 TTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCA 203
           || ||  ||| ||      ||| |   |||   | |    || |  |   ||||    |
Db    6888 TTGTAGTGGTAGTGGGAGTAGTGGTTGTAGTGGTTGTAGTGGTAGTGGGAGTAGTGGTTG 6829

Qy     204 TCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAGTAGTCGAGTTGGG 255
           | ||| |   ||| |  |||| | | | ||||| | | ||||| |   ||||
Db    6828 TAGTGGTAGTGGGAGTAGTGGTTGTAGTGGTAGTGGTGGTAGTGGTAGTGGG 6777


RESULT 14
US-08-663-566A-8
; Sequence 8, Application US/08663566A
; Patent No. 5853733
;  GENERAL INFORMATION:
;    APPLICANT:  Cochran, Mark D
;    APPLICANT:  Macdonald, Richard D
;    TITLE OF INVENTION:  Recombinant Herpesvirus of Turkeys
;    TITLE OF INVENTION:  and Uses Thereof
;    NUMBER OF SEQUENCES:  56
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  John P. White
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
```

```
;       COUNTRY:  USA
;       ZIP:  10036
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/663,566A
;       FILING DATE:  June 13, 1996
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  White, John P
;       REGISTRATION NUMBER:  28,678
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212)278-0400
;       TELEFAX:  (212)391-0526
;       TELEX:  422523
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1506 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  1..1506
US-08-663-566A-8

  Query Match             8.0%;  Score 28.6;  DB 2;  Length 1506;
  Best Local Similarity   61.3%;  Pred. No. 9.7;
  Matches   46;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy   190 AATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAGTAGTCGA 249
         || || ||||  ||| ||  ||||||  | |||  | | ||||  | ||  || |  | |
Db  1383 AAGAGGAACACCCATGGTTATTACGGTTACGGCAGTTTTGGGATTGGCTGTAATTTTAGG 1442

Qy   250 GTTGGGCAGCAACTT 264
         | ||||  |  | | |
Db  1443 GATGGGGATAATCAT 1457


RESULT 15
US-08-023-610-8
; Sequence 8, Application US/08023610
; Patent No. 5928648
;  GENERAL INFORMATION:
;    APPLICANT:  Cochran Ph.D, Mark D
;    APPLICANT:  Macdonald Ph.D., Richard D
;    TITLE OF INVENTION:  Recombinant Herpesvirus of Turkeys
;    TITLE OF INVENTION:  and Uses Thereof
;    NUMBER OF SEQUENCES:  47
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  John P. White
;      STREET:  30 Rockefeller Plaza
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10112
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/023,610
;      FILING DATE:  February 26, 1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White Esq, John P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)977-9550
;      TELEFAX:  (212)664-0525
;      TELEX:  422523
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1506 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..1506
US-08-023-610-8

  Query Match             8.0%;  Score 28.6;  DB 2;  Length 1506;
  Best Local Similarity   61.3%;  Pred. No. 9.7;
  Matches   46;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy   190 AATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAGTAGTCGA 249
         || || ||||  ||| ||  ||||||  | |||  | | ||||  | ||  || |  | |
Db  1383 AAGAGGAACACCCATGGTTATTACGGTTACGGCAGTTTTGGGATTGGCTGTAATTTTAGG 1442

Qy   250 GTTGGGCAGCAACTT 264
         | ||||  |  | | |
Db  1443 GATGGGGATAATCAT 1457


Search completed: March 11, 2003, 15:26:16
Job time : 75.5 secs
```

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: March 11, 2003, 13:59:58 ; Search time 743 Seconds
(without alignments)
320.925 Million cell updates/sec

Title: US-09-936-011-3
Perfect score: 357
Sequence: 1 atgtctagggaaataaccgc.....gtaatggtgcttgtggttaa 357

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 478924 seqs, 333959956 residues

Total number of hits satisfying chosen parameters: 957848

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		36.6	10.3	4039	10	US-09-925-301-371	Sequence 371, App
2	c	34.2	9.6	493	10	US-09-864-761-5863	Sequence 5863, Ap
3	c	32.8	9.2	2361	10	US-09-815-242-8565	Sequence 8565, Ap
4	c	32.2	9.0	234	10	US-09-864-761-28286	Sequence 28286, A
5		31.6	8.9	155074	9	US-10-026-188-6	Sequence 6, Appli
6	c	31.2	8.7	472	10	US-09-864-761-11710	Sequence 11710, A
7	c	30.8	8.6	276	10	US-09-864-761-25120	Sequence 25120, A
8	c	30.8	8.6	537	10	US-09-864-761-8390	Sequence 8390, Ap
9		30.4	8.5	867	10	US-09-737-178-97	Sequence 97, Appl
10		30.4	8.5	967	10	US-09-983-531A-9	Sequence 9, Appli
11	c	30.4	8.5	1333	10	US-09-737-178-130	Sequence 130, App
12		30.4	8.5	1668	10	US-09-792-200B-3	Sequence 3, Appli
13	c	30.4	8.5	127197	9	US-09-754-853A-1	Sequence 1, Appli
14	c	30.2	8.5	574	10	US-09-864-761-228	Sequence 228, App
15	c	30.2	8.5	669	10	US-09-864-761-17051	Sequence 17051, A
16	c	30.2	8.5	759	10	US-09-966-881-33	Sequence 33, Appl
17		30	8.4	401	10	US-09-864-761-3936	Sequence 3936, Ap
18		30	8.4	446	10	US-09-864-761-20699	Sequence 20699, A
19		29.6	8.3	287	9	US-09-535-459-1849	Sequence 1849, Ap
20	c	29.6	8.3	2960	9	US-10-098-841-16	Sequence 16, Appl
21	c	29.6	8.3	3098	9	US-10-098-841-15	Sequence 15, Appl
22		29.6	8.3	4285	9	US-10-104-580-1	Sequence 1, Appli
23	c	29.6	8.3	9567	12	US-10-016-768-12	Sequence 12, Appl
24	c	29.6	8.3	9574	12	US-10-016-768-9	Sequence 9, Appli
25		29.4	8.2	594	10	US-09-864-761-7908	Sequence 7908, Ap
26		29.4	8.2	730	10	US-09-764-864-248	Sequence 248, App
27		29.4	8.2	753	10	US-09-764-853-177	Sequence 177, App
28		29.4	8.2	818	10	US-09-764-864-676	Sequence 676, App
29		29.4	8.2	873	10	US-09-764-853-178	Sequence 178, App
30		29.4	8.2	1030	10	US-09-878-574-4306	Sequence 4306, Ap
31	c	29.4	8.2	2901	10	US-09-801-368-371	Sequence 371, App
32	c	29.4	8.2	2934	9	US-09-938-901-7	Sequence 7, Appli
33	c	29.4	8.2	10236	10	US-09-764-853-878	Sequence 878, App
34		29.2	8.2	456	10	US-09-864-761-5185	Sequence 5185, Ap
35	c	29.2	8.2	474	10	US-09-864-761-1364	Sequence 1364, Ap
36		29.2	8.2	887	10	US-09-864-761-21484	Sequence 21484, A
37		29.2	8.2	1075	10	US-09-864-761-19241	Sequence 19241, A
38	c	29.2	8.2	1249	10	US-09-881-752A-223	Sequence 223, App
39		29.2	8.2	1403	10	US-09-864-761-2513	Sequence 2513, Ap
40	c	29.2	8.2	1635	10	US-09-864-761-20241	Sequence 20241, A
41	c	29.2	8.2	1973	10	US-09-864-761-3471	Sequence 3471, Ap
42		29.2	8.2	1981	10	US-09-864-761-4745	Sequence 4745, Ap
43	c	29	8.1	370	10	US-09-770-791-356	Sequence 356, App
44		29	8.1	479	10	US-09-864-761-768	Sequence 768, App
45		29	8.1	496	9	US-10-040-739-276	Sequence 276, App

ALIGNMENTS

RESULT 1
US-09-925-301-371
; Sequence 371, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 371
; LENGTH: 4039
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1085)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-371

Query Match 10.3%; Score 36.6; DB 10; Length 4039;
Best Local Similarity 49.7%; Pred. No. 0.25;
Matches 93; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

```
Qy    123 TGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGC 182
          |||   |   | ||||||   | |   ||  |  |  | || | |  | | ||  ||
Db    586 TGACCCCCTGGCACCAACACCCGCCAGTACTGTGACTTCCAAAGCCAGAGCCACATGTGC 645

Qy    183 TGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAG 242
          | ||   | | |  || ||    ||    |   || |||| ||||||||     |
Db    646 TCATCAAACTTGCATTAAGCAGTTGGCGGGAGATGGCTGTGGAGCTGGGGGTTTAAGTGA 705

Qy    243 TAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCACT 302
          | ||    ||  ||  |  ||| |   | ||    |  | ||| | |||| |||||   |
Db    706 TGGTTCTCTTTTGCTCCCTCTTYTGAGGGTAAAGCTACTGTCTTTCTTAAGAGTGTATTT 765
```

```
Qy     303 ATTACAA 309
           ||  |||
Db     766 ATGCCAA 772
```

RESULT 2
US-09-864-761-5863/c
; Sequence 5863, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 5863
; LENGTH: 493
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004169.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
US-09-864-761-5863

Query Match 9.6%; Score 34.2; DB 10; Length 493;
Best Local Similarity 49.8%; Pred. No. 0.53;
Matches 115; Conservative 0; Mismatches 113; Indels 3; Gaps 1;

```
Qy      30 TAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTT 89
           | | ||||    |  || || ||||| | | || ||||   |  |  || |  ||| |
Db     424 TGACAATGGTGGTGGTGATGGTGGTGATAGGGTGGTGGTGATGGTGATAGTAGTGGTGGG 365

Qy      90 CATGCAGCAAGCAGCTGCGACACATTCTGGGGG---TGATTACGGAGTCCCAACATTTTC 146
            ||| |    |  | ||       ||  ||| ||   || | | || |     |    |
Db     364 AATGGATGGTGGTGGTGATGGTGATGGTGGTGGTGTTGGTGATGGTGGTGATAGTGCTGG 305

Qy     147 TAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCG 206
           | | | ||||     | | || |||  | |    | || |  | || || | |||  | |
Db     304 TGAGGATGGTGATGGTGGTGATGGTGGAGGTGATGGTGGTAGTTATGGTGATAATGGTGG 245

Qy     207 TGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAGTAGTCGAGTTGGGCA 257
           |  |   |   | ||||   ||| | | | | |  ||||   | |||| |
Db     244 TAGTGGTGATGGTGGTGGTGTTGGTGATGGTGGTGATAGTATTGGTGGGAA 194
```

RESULT 3
US-09-815-242-8565/c
; Sequence 8565, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8565
; LENGTH: 2361
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2361)
US-09-815-242-8565

Query Match 9.2%; Score 32.8; DB 10; Length 2361;
Best Local Similarity 58.0%; Pred. No. 2.9;
Matches 58; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```
Qy      28 AATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCG 87
           |||  ||||||  || |    |||||| |||||||| |||||  | |||||      ||
Db     419 AATTTGAATGTTGCTGTGCGAGTTGGTTTTTGTGTTGAGGCTTGTTGTGTATCTACATCG 360

Qy      88 TTCATGCAGCAAGCAGCTGCGACACATTCTGGGGGTGATT 127
             | |   |    | | |||   |  | |||  | || ||
```

```
Db    359 CGCGTTTCGTTTACTGTTGCTGAAACTGCTGTAGATGGTT 320


RESULT 4
US-09-864-761-28286/c
; Sequence 28286, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28286
;   LENGTH: 234
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC004547.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.94
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.77
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.75
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.86
;   OTHER INFORMATION: NT HIT: Y11842.1, EVALUE 1.20e-01
;   OTHER INFORMATION: EST_HUMAN HIT: BE143078.1, EVALUE 3.60e-02
US-09-864-761-28286

  Query Match             9.0%;  Score 32.2;  DB 10;  Length 234;
  Best Local Similarity   46.9%;  Pred. No. 1.6;
  Matches  100;  Conservative    0;  Mismatches  113;  Indels    0;  Gaps    0;

Qy      27 CAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGC 86
           || | |   ||| | ||||       ||||  ||     || |   |||||  |     |
Db     221 CACTTACTTTGTACTTATTTAGAACAGTGTCTGGTACATGACAAGCTTTGGCTATGATGA 162

Qy      87 GTTCATGCAGCAAGCAGCTGCGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTC 146
             | || |||  | ||      ||  ||   || | | |       |     |
Db     161 TGAGTAGGAGAAAGATGATGATGGTGATAGTGATAGTAGTGAAGATAATGATAGTGGTGG 102

Qy     147 TAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCG 206
           | | ||| |||  |||   |||| |    |    | |||| || || || ||| | ||
Db     101 TGATGGTAGTAGTTATGATGACGATGGTGGTGATGATGATATTGATGGTGACAGTGATAA 42

Qy     207 TGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGT 239
           || |   ||  | ||||    ||| | | | ||
Db      41 TGGTGGTGGTAGTGGTGATGATGGTGATGGTGT 9


RESULT 5
US-10-026-188-6
; Sequence 6, Application US/10026188
; Patent No. US20020164645A1
; GENERAL INFORMATION:
;  APPLICANT: Zuker, Charles S.
;  APPLICANT:  Zhang, Yifeng
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Assays for Taste Receptor Cell Specific
;  TITLE OF INVENTION:  Ion Channel
;  FILE REFERENCE: 02307E-114910US
;  CURRENT APPLICATION NUMBER: US/10/026,188
;  CURRENT FILING DATE:  2001-12-21
;  PRIOR APPLICATION NUMBER: US 60/259,379
;  PRIOR FILING DATE: 2000-12-29
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
;   LENGTH: 155074
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: human genomic region containing ltrpc5
;   OTHER INFORMATION: (Human Chromosome 11p15.5 PAC clone pDJ915f1
;   OTHER INFORMATION: containing KvLQT1 gene)
US-10-026-188-6

  Query Match             8.9%;  Score 31.6;  DB 9;  Length 155074;
  Best Local Similarity   45.0%;  Pred. No. 46;
  Matches  118;  Conservative    0;  Mismatches  144;  Indels    0;  Gaps    0;

Qy      45 TGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGC 104
           ||  || | ||   | | ||||  ||  ||||  || |||   | |||      | |
Db   44159 TGGGGTAGTGGTGGTTATGGTGGTGTTGGTTGTTATGGTGGTAGTGATGATCATGGTGGT 44218

Qy     105 TGCGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAG 164
             | |||          ||||||   || || |       |  | | | || ||    | |
Db   44219 GGTGACTATAGTCACTGGTGATGCTGGTGTTCATGGTGATGGTGATGATGATAGTGGTGG 44278

Qy     165 AGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGG 224
           |   ||   ||  | |||| |  |  | ||    || ||   | | | |    | ||
Db   44279 TGGTTGTGGCAGTGCTGCTGCTGATGGTGGTGGTGATGATGAGGGTGATGATGACAATGA 44338

Qy     225 GTCTGGGGGTAGCGTTAGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTC 284
           || ||  | | | | | ||  |     |||       |  |  || | |   || ||
Db   44339 GTATGATGATGGTGGTGGTGATGATCGTGGTTGTGGTGTGTTGATGGTGGTGTTTATTGT 44398

Qy     285 TGTGTTAATAGTGTCACTATTA 306
           | | | ||  ||  | | |||
Db   44399 GGCGGTGATGATGATAGTGTTA 44420


RESULT 6
US-09-864-761-11710/c
```

```
; Sequence 11710, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 11710
;   LENGTH: 472
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC004547.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.94
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.77
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.75
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.86
US-09-864-761-11710

  Query Match             8.7%;  Score 31.2;  DB 10;  Length 472;
  Best Local Similarity   47.1%;  Pred. No. 4.4;
  Matches   96;  Conservative    0;  Mismatches  108;  Indels    0;  Gaps    0;

Qy      36 TGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCA 95
           ||| | ||||       |||| ||     || |   |||||  |      |       | |
Db     470 TGTACTTATTTAGAACAGTGTCTGGTACATGACAAGCTTTGGCTATGATGATGAGTAGGA 411

Qy      96 GCAAGCAGCTGCGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACGGTGG 155
```

```
           | |||  | ||      ||  ||   ||  | | |        |    |  | | ||| |
Db     410 GAAAGATGATGATGGTGATAGTGATAGTAGTGAAGATAATGATAGTGGTGGTGATGGTAG 351

Qy     156 TATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCGTGCTTACGG 215
           ||  |||   |||| |     |    | |||| || || || ||| | ||  || |   ||
Db     350 TAGTTATGATGACGATGGTGGTGATGATGATATTGATGGTGACAGTGATAATGGTGGTGG 291

Qy     216 GTGCGGTGGGTCTGGGGGTAGCGT 239
             | ||||    ||| | | | ||
Db     290 TAGTGGTGATGATGGTGATGGTGT 267


RESULT 7
US-09-864-761-25120/c
; Sequence 25120, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 25120
;   LENGTH: 276
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC012153.10
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.1
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.8
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
```

```
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL =   3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1
US-09-864-761-25120

  Query Match             8.6%;  Score 30.8;  DB 10;  Length 276;
  Best Local Similarity   49.1%;  Pred. No. 4.6;
  Matches  109;  Conservative    0;  Mismatches  112;  Indels    1;  Gaps    1;

Qy      32 AGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCA 91
           ||  ||    || || || ||||||    || |||    |  | ||  || ||  | | |
Db     267 AGTGTGATGGTAGTGGTGATGGTGTGATGGTGGTGATGGTGATGGTGGTGGTGATGGTGA 208

Qy      92 TGCAGCAAGCAGCTGCGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACG 151
           || |    |  | ||       ||  ||| |||| | | || |          |  | |
Db     207 TG-ATGGTGGTGGTGATGGTGATGATGGTGGTGGTGATGGTGGTGATGATGGTGGTGATT 149

Qy     152 GTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCGTGCTT 211
           || ||   | | | || |||    |    | ||||  |  | || ||  | || ||| |
Db     148 GTAGTGATTGTGGTGATGGTGATGGTGGTGATGATGGTGGTGGTGACGGTGATGGTGGTG 89

Qy     212 ACGGGTGCGGTGGGTCTGGGGGTAGCGTTAGTAGTCGAGTTG 253
           ||||  | | | |   ||  |||   | | ||  | | | ||
Db      88 ACGGTAGTGATAGTGGTGATGGTGATGATGGTGATGGTGGTG 47


RESULT 8
US-09-864-761-8390/c
; Sequence 8390, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 8390
;   LENGTH: 537
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC012153.10
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.1
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.8
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL =   3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1
US-09-864-761-8390

  Query Match             8.6%;  Score 30.8;  DB 10;  Length 537;
  Best Local Similarity   49.1%;  Pred. No. 6.2;
  Matches  109;  Conservative    0;  Mismatches  112;  Indels    1;  Gaps    1;

Qy      32 AGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCA 91
           ||  ||    || || || ||||||    || |||    |  | ||  || ||  | | |
Db     291 AGTGTGATGGTAGTGGTGATGGTGTGATGGTGGTGATGGTGATGGTGGTGGTGATGGTGA 232

Qy      92 TGCAGCAAGCAGCTGCGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACG 151
           || |    |  | ||       ||  ||| |||| | | || |          |  | |
Db     231 TG-ATGGTGGTGGTGATGGTGATGATGGTGGTGGTGATGGTGGTGATGATGGTGGTGATT 173

Qy     152 GTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCGTGCTT 211
           || ||   | | | || |||    |    | ||||  |  | || ||  | || ||| |
Db     172 GTAGTGATTGTGGTGATGGTGATGGTGGTGATGATGGTGGTGGTGACGGTGATGGTGGTG 113

Qy     212 ACGGGTGCGGTGGGTCTGGGGGTAGCGTTAGTAGTCGAGTTG 253
           ||||  | | | |   ||  |||   | | ||  | | | ||
Db     112 ACGGTAGTGATAGTGGTGATGGTGATGATGGTGATGGTGGTG 71


RESULT 9
US-09-737-178-97
; Sequence 97, Application US/09737178
; Patent No. US20010029295A1
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Homer, Mary
;  APPLICANT:  Secrist, Heather
;  TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THE DIAGNOSIS
;  TITLE OF INVENTION:  AND TREATMENT OF B. MICROTI INFECTION
;  FILE REFERENCE:  210121.426C9
;  CURRENT APPLICATION NUMBER: US/09/737,178
;  CURRENT FILING DATE:  2000-12-13
;  NUMBER OF SEQ ID NOS:  144
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 97
;   LENGTH: 867
;   TYPE: DNA
;   ORGANISM: Babesia microti
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(867)
;   OTHER INFORMATION: n=A,T,C or G
US-09-737-178-97
```

```
  Query Match             8.5%;  Score 30.4;  DB 10;  Length 867;
  Best Local Similarity   49.4%;  Pred. No. 10;
  Matches   79;  Conservative    0;  Mismatches   81;  Indels    0;  Gaps    0;

Qy      32 AGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCA 91
           | | |||   | | |  | || |||| ||| || ||| |  |||   ||   |   |
Db     650 AAATTGTTGATGTGGCAGCTGTTGTTGATGTGGTAGCTGTTGTTGATGTGGCAGTTGTTG 709

Qy      92 TGCAGCAAGCAGCTGCGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACG 151
           |  |    |||| ||      ||   |  | || |   |  ||  ||    ||  | | |
Db     710 TTGATGTGGCAGTTGTTGTTGATGTGGTAGCTGTTGTTGATGTAGCAGATGTTGTTGATG 769

Qy     152 GTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAA 191
           |   || |||  ||| ||||| |  ||  ||  | |||
Db     770 TAGCAGTACATACTGACAGTACATGTGCATGTGTGTGTAA 809


RESULT 10
US-09-983-531A-9
; Sequence 9, Application US/09983531A
; Patent No. US20020147132A1
; GENERAL INFORMATION:
;  APPLICANT: Fujisawa, Atsuko
;  APPLICANT:  Yamakawa, Toru
;  APPLICANT:  Shirakawa, Kamon
;  APPLICANT:  Chitose, Orii
;  APPLICANT:  Ogawa, Naoki
;  TITLE OF INVENTION: Meltrins
;  FILE REFERENCE: 11-22-99 sequence submission
;  CURRENT APPLICATION NUMBER: US/09/983,531A
;  CURRENT FILING DATE:  2001-10-24
;  PRIOR APPLICATION NUMBER: JP 8-61756
;  PRIOR FILING DATE: 1996-02-23
;  PRIOR APPLICATION NUMBER: PCT/JP96/03017
;  PRIOR FILING DATE: 1996-10-17
;  NUMBER OF SEQ ID NOS: 44
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
;   LENGTH: 967
;   TYPE: DNA
;   ORGANISM: Unknown Organism
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism:  Clone:
;   OTHER INFORMATION: JM109(pBShuM-gamma-G238), human meltrin gamma
;   NAME/KEY: CDS
;   LOCATION: (1)..(966)
US-09-983-531A-9

  Query Match             8.5%;  Score 30.4;  DB 10;  Length 967;
  Best Local Similarity   52.3%;  Pred. No. 11;
  Matches   67;  Conservative    0;  Mismatches   61;  Indels    0;  Gaps    0;

Qy      83 TGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGGGGTGATTACGGAGTCCCAACAT 142
           |||  |||  ||    | ||| ||   || || |||||   | | ||| || | ||| ||
Db     710 TGGATTTCCAGCTAGGATCAGATGTTCCAGATCCTGGGATGGTTAACGAAGGCACAAAAT 769

Qy     143 TTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATC 202
           |  |   ||    || | |||| ||       |   || || || |  |   ||  ||
Db     770 GTGGTGCTGGAAAGATCTGTAGAAACTTCCAGTGTGTAGATGCTTCTGTTCTGAATTATG 829

Qy     203 ATCGTGCT 210
           |  ||| |
Db     830 ACTGTGAT 837


RESULT 11
US-09-737-178-130/c
; Sequence 130, Application US/09737178
; Patent No. US20010029295A1
; GENERAL INFORMATION:
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Lodes, Michael J.
;  APPLICANT:  Houghton, Raymond L.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Homer, Mary
;  APPLICANT:  Secrist, Heather
;  TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THE DIAGNOSIS
;  TITLE OF INVENTION:  AND TREATMENT OF B. MICROTI INFECTION
;  FILE REFERENCE:  210121.426C9
;  CURRENT APPLICATION NUMBER: US/09/737,178
;  CURRENT FILING DATE:  2000-12-13
;  NUMBER OF SEQ ID NOS:  144
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 130
;   LENGTH: 1333
;   TYPE: DNA
;   ORGANISM: Babesia microti
US-09-737-178-130

  Query Match             8.5%;  Score 30.4;  DB 10;  Length 1333;
  Best Local Similarity   49.4%;  Pred. No. 13;
  Matches   79;  Conservative    0;  Mismatches   81;  Indels    0;  Gaps    0;

Qy      32 AGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCA 91
           | | |||   | | |  | || |||| ||| || ||| |  |||   ||   |   |
Db     684 AAATTGTTGATGTGGCAGCTGTTGTTGATGTGGTAGCTGTTGTTGATGTGGCAGTTGTTG 625

Qy      92 TGCAGCAAGCAGCTGCGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACG 151
           |  |    |||| ||      ||   |  | || |   |  ||  ||    ||  | | |
Db     624 TTGATGTGGCAGTTGTTGTTGATGTGGTAGCTGTTGTTGATGTAGCAGATGTTGTTGATG 565

Qy     152 GTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAA 191
           |   || |||  ||| ||||| |  ||  ||  | |||
Db     564 TAGCAGTACATACTGACAGTACATGTGCATGTGTGTGTAA 525


RESULT 12
US-09-792-200B-3
; Sequence 3, Application US/09792200B
; Patent No. US20020042368A1
; GENERAL INFORMATION:
;  APPLICANT: Immunex Corporation
;  APPLICANT:  Fanslow, William C.
;  APPLICANT:  Poindexter, Kurt
;  APPLICANT:  Cerretti, Douglas P.
;  APPLICANT:  Black, Roy A.
;  TITLE OF INVENTION: INTEGRIN ANTAGONISTS
;  FILE REFERENCE: 2958-A
;  CURRENT APPLICATION NUMBER: US/09/792,200B
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: US 60/184,865
;  PRIOR FILING DATE: 2000-02-25
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
;   LENGTH: 1668
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Fusion Construct
;   NAME/KEY: CDS
;   LOCATION: (46)..(1647)
;   OTHER INFORMATION:
US-09-792-200B-3

  Query Match             8.5%;  Score 30.4;  DB 10;  Length 1668;
  Best Local Similarity   52.3%;  Pred. No. 14;
  Matches   67;  Conservative    0;  Mismatches   61;  Indels    0;  Gaps    0;

Qy      83 TGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGGGGTGATTACGGAGTCCCAACAT 142
           |||  |||  ||    | ||| ||   || || |||||   | | ||| || | ||| ||
Db     671 TGGATTTCCAGCTAGGATCAGATGTTCCAGATCCTGGGATGGTTAACGAAGGCACAAAAT 730

Qy     143 TTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATC 202
           |  |   ||    || | |||| ||       |   || || || |  |.  ||  ||
```

```
Db    731 GTGGTGCTGGAAAGATCTGTAGAAACTTCCAGTGTGTAGATGCTTCTGTTCTGAATTATG 790

Qy    203 ATCGTGCT 210
          |  ||| |
Db    791 ACTGTGAT 798


RESULT 13
US-09-754-853A-1/c
; Sequence 1, Application US/09754853A
; Publication No. US20030005491A1
; GENERAL INFORMATION:
;  APPLICANT: Hauge, Brian M.
;  APPLICANT:  Parnell, Laurence D.
;  APPLICANT:  Parsons, Jeremy D.
;  APPLICANT:  Wang, Ming Li
;  TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;  TITLE OF INVENTION:  Soybean Cyst Nematode Resistance
;  FILE REFERENCE: 38-10(15810)B
;  CURRENT APPLICATION NUMBER: US/09/754,853A
;  CURRENT FILING DATE:  2001-01-05
;  PRIOR APPLICATION NUMBER: US 60/174,880
;  PRIOR FILING DATE: 2000-01-07
;  NUMBER OF SEQ ID NOS: 1119
; SEQ ID NO 1
;   LENGTH: 127197
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: 515002_region_G2
US-09-754-853A-1

  Query Match             8.5%;  Score 30.4;  DB 9;  Length 127197;
  Best Local Similarity   59.1%;  Pred. No. 99;
  Matches   52;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy  5 CTAGGGAAATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTG 64
        |||  ||||||    |||||     |||| ||  |||| ||||| | || |    |||||
Db  24247 CTATTGAAATATTATCTCGATTATGTAAGGATCAGCCTTTTGTTCTGGGAGACACTGTTG 24188

Qy     65 TGGCTTTCTTTGTATTGCTGGCGTTCAT 92
          |  ||  || || ||   |  | |  ||
Db  24187 TTACTGCCTCTGGATGTATATCTTCTAT 24160


RESULT 14
US-09-864-761-228/c
; Sequence 228, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 228
;   LENGTH: 574
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL079301.14
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 7.8
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.2
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.2
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 8.4
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.7
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.3
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.6
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 7.8
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL =   7
US-09-864-761-228

  Query Match             8.5%;  Score 30.2;  DB 10;  Length 574;
  Best Local Similarity   55.1%;  Pred. No. 9.8;
  Matches   59;  Conservative    0;  Mismatches   48;  Indels    0;  Gaps    0;

Qy  147 TAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCG 206
        | | | ||||    |||| || |||   |     | | ||| | || |||    |  | |
Db  367 TGAGGATGGTTGTCATAGTGATGGTGATAATGGTGATCATTGTGATGGTAGTGGTGGTGG 308

Qy  207 TGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAGTAGTCGAGTTG 253
        |||| | ||    |||||  |||  | | | | | || |||  | ||
Db  307 TGCTGATGGTGATGGTGGTGCTGCTGCTGGTGATGGTTGTCATGGTG 261


RESULT 15
US-09-864-761-17051/c
; Sequence 17051, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
```

```
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 17051
;   LENGTH: 669
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL079301.14
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 7.8
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.2
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.2
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 8.4
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.7
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.3
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.6
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 7.8
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL =   7
;   OTHER INFORMATION: EST_HUMAN HIT: BE011604.1, EVALUE 1.10e-01
;   OTHER INFORMATION: NT HIT: gi9628932, EVALUE 1.10e-01
US-09-864-761-17051

  Query Match             8.5%;  Score 30.2;  DB 10;  Length 669;
  Best Local Similarity   55.1%;  Pred. No. 11;
  Matches   59;  Conservative    0;  Mismatches   48;  Indels    0;  Gaps    0;

Qy    147 TAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCG 206
          | | | ||||    |||| || |||   |     | | ||| | || |||    |  | |
Db    301 TGAGGATGGTTGTCATAGTGATGGTGATAATGGTGATCATTGTGATGGTAGTGGTGGTGG 242

Qy    207 TGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAGTAGTCGAGTTG 253
          |||| | ||    ||||| |||  | | | | | || |||  | ||
Db    241 TGCTGATGGTGATGGTGGTGCTGCTGCTGGTGATGGTTGTCATGGTG 195


Search completed: March 12, 2003, 08:12:50
Job time : 847 secs
```

OM nucleic - nucleic search, using sw model

Run on: March 12, 2003, 06:32:01 ; Search time 1191.5 Seconds
(without alignments)
4852.529 Million cell updates/sec

Title: US-09-936-011-3
Perfect score: 357
Sequence: 1 atgtctagggaaataaccgc.....gtaatggtgcttgtggttaa 357

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	43.8	12.3	1101	17	CNS00LT2	AL078714 Drosophil
2	42.8	12.0	1080	17	CNS00EPP	AL069494 Drosophil
3	40.6	11.4	902	17	CNS006QP	AL065804 Drosophil
4	39.2	11.0	598	13	BJ032222	BJ032222 BJ032222
5	38.4	10.8	1101	17	CNS00FOO	AL070854 Drosophil
6	37.8	10.6	884	17	CNS006U0	AL065923 Drosophil
7	37.6	10.5	1101	17	CNS0100X	AL098379 Drosophil
c 8	37.4	10.5	401	10	BB671302	BB671302 BB671302
9	37.2	10.4	474	12	BF417205	BF417205 UI-R-CN0-
10	36.6	10.3	873	14	BQ215809	BQ215809 AGENCOURT
11	36.4	10.2	486	9	AA134045	AA134045 zo17d11.r
12	36.2	10.1	922	17	CNS0073W	AL066784 Drosophil
13	36	10.1	328	10	BB224881	BB224881 BB224881
c 14	35.8	10.0	498	17	AQ927406	AQ927406 RPCI-23-2
c 15	35.8	10.0	939	17	CNS00CNG	AL059400 Drosophil
c 16	35.6	10.0	1207	17	CNS015ZN	AL106109 Drosophil
c 17	35.4	9.9	644	13	BM605179	BM605179 170006870
18	35.2	9.9	554	17	AQ523013	AQ523013 HS_5230_A
c 19	35.2	9.9	572	17	CNS043SN	AL273200 Tetraodon
c 20	35.2	9.9	770	12	BF678875	BF678875 602153190
c 21	35.2	9.9	1013	17	CNS016KT	AL106871 Drosophil
c 22	35.2	9.9	1101	17	CNS00LO0	AL068607 Drosophil
c 23	34.8	9.7	723	17	AZ627893	AZ627893 1M0469I18
24	34.6	9.7	477	14	BQ353390	BQ353390 PM1-HT045
25	34.6	9.7	894	12	BG776437	BG776437 602663521
c 26	34.4	9.6	563	13	BJ336822	BJ336822 BJ336822
27	34.4	9.6	721	13	BJ137716	BJ137716 BJ137716
28	34.2	9.6	330	9	AI501178	AI501178 UI-R-C2p-
c 29	34.2	9.6	336	10	AV417568	AV417568 AV417568
c 30	34.2	9.6	349	10	AV413481	AV413481 AV413481
31	34.2	9.6	363	12	BF418198	BF418198 UI-R-CN0-
32	34.2	9.6	390	10	AW524847	AW524847 UI-R-BO0-
33	34.2	9.6	391	9	AA900710	AA900710 UI-R-E0-b
c 34	34.2	9.6	394	10	AV423714	AV423714 AV423714
c 35	34.2	9.6	410	10	AV408287	AV408287 AV408287
c 36	34.2	9.6	424	10	AV428494	AV428494 AV428494
c 37	34.2	9.6	427	10	AV408779	AV408779 AV408779
c 38	34.2	9.6	429	10	AV412428	AV412428 AV412428
39	34.2	9.6	452	9	AI705708	AI705708 UI-R-AD1-
40	34.2	9.6	452	12	BF417072	BF417072 UI-R-CN0-
41	34.2	9.6	508	10	BE109288	BE109288 UI-R-BJ1-
c 42	34.2	9.6	543	17	CNS00VWP	AL092839 Arabidops
c 43	34.2	9.6	566	17	AZ276668	AZ276668 RPCI-23-1
44	34.2	9.6	585	13	BM385308	BM385308 UI-R-EB1-
45	34.2	9.6	602	13	BM391584	BM391584 UI-R-DY0-

ALIGNMENTS

RESULT 1
CNS00LT2
LOCUS CNS00LT2 1101 bp DNA linear GSS 14-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC: BACR48P19 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL078714
VERSION AL078714.1 GI:5102004
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (11-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial

EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

```
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACR48P19"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
BASE COUNT      469 a      6 c     69 g    151 t    406 others
ORIGIN
```

```
  Query Match             12.3%;  Score 43.8;  DB 17;  Length 1101;
  Best Local Similarity   13.7%;  Pred. No. 0.62;
  Matches   37;  Conservative  128;  Mismatches  106;  Indels    0;  Gaps    0;

Qy      28 AATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCG 87
           :: || ||  ::  :  :::::|:::| ::::::::::|| :::   ::: :   ::: :
Db     831 WDAAAAAAAAKKTKDKGKKKKKTKKKTTKKKKKKKKKKGGGKKDAAAKKKKKGTKKKKGK 890

Qy      88 TTCATGCAGCAAGCAGCTGCGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCT 147
           :   :|  :   :  : ::  | |  :  ::::|: |  :  :| ::        ::|: |
Db     891 KGKGKGGGKGKGKGGKKDDDAAAAKKKGKKKKKGKGGGGKKKKGGKKGKGKGGKKKTKKT 950

Qy     148 AACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCGT 207
             ::::::|:|       ::   :| ||::| |  :: :::: :: :|     : |   |
Db     951 KKKKKKKKKADAAKGKTKKRAAADAAAADWTDAATKKKKKKKKDKKKTKTKKKDAAAAAT 1010

Qy     208 GCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAGTAGTCGAGTTGGGCAGCAACTTATT 267
           : ::  :::|: ::|:::: ::::: : : :|| :::|: : |: :|:  :    |: ::
Db    1011 KKKKKKKKKTKKKKTKKKKKKKKKDADKKKKTTGKKDGKGKGGKGKGKKGKKKTKTKKKK 1070

Qy     268 GTGTTAGCTATTGTTTCTGTGTTAATAGTGT 298
           :::::   : || ::  ::::: |:: ::::
Db    1071 KKKKKKTKKKTTADKGKKKKKKAADDKKKKK 1101
```

```
RESULT 2
CNS00EPP
LOCUS       CNS00EPP                1080 bp    DNA     linear   GSS 04-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC:
            BACR29N07 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL069494
VERSION     AL069494.1  GI:4949637
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1080)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .1080
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACR29N07"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
BASE COUNT      171 a    105 c    219 g    225 t    360 others
ORIGIN
```

```
  Query Match             12.0%;  Score 42.8;  DB 17;  Length 1080;
  Best Local Similarity   20.7%;  Pred. No. 1.1;
  Matches   51;  Conservative  102;  Mismatches   93;  Indels    0;  Gaps    0;

Qy      45 TGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGC 104
           ::::::|::: ::   ::: ::: ::| :|  : :|: :|: :     : |: |   ::
Db     627 WDKWKKTKKKTKWGTRKWWTKKRGKDTRKTGTKTKTKKDGRKKGGKGRDTADGAKTDRKK 686

Qy     105 TGCGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAG 164
           |: |:   : :  |||:|:| ||  :|:|:   :| :|:|: :::   :|:  :  :::|
Db     687 TRKGDTKKRRDRAGGGKGDGGTTTDKGDGDWWADATDTWTKKKDDTTWWGDATDKRWWWG 746

Qy     165 AGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGG 224
           ::: :::: ::    |   :|:: |:|:|: :| :|: ::     |::  ::|: ::|:
Db     747 WRWAKDDDDWWAGRGAAGARADDRTDADARGWAARAKKDKTTAWATDWTTRKTRTDDTDT 806

Qy     225 GTCTGGGGGTAGCGTTAGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTC 284
           |:  :|:|: :| ::: | ||| :::  :|:  : || : : :    :   : :| |||
Db     807 GDGRKGRGKGWGKRKDGGGAGTRKRKGARGKGTDAAAAWAWGDTRTADKTADTKTTTTTA 866

Qy     285 TGTGTT 290
           | |:|:
Db     867 TTTKTK 872
```

```
RESULT 3
CNS006QP
LOCUS       CNS006QP                902 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR14J09 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL065804
VERSION     AL065804.1  GI:4944772
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 902)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
```

```
FEATURES             Location/Qualifiers
     source          1. .902
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACR14J09"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
BASE COUNT      298 a     50 c     61 g    142 t    351 others
ORIGIN

  Query Match           11.4%;  Score 40.6;  DB 17;  Length 902;
  Best Local Similarity  22.4%;  Pred. No. 4;
  Matches   57;  Conservative   93;  Mismatches  105;  Indels    0;  Gaps    0;

Qy      36 TGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCA 95
           |:::  : :: :||:|::| :||:: || ::  |:: :|::| |:| ::| |::  ::
Db     524 TKKKTKKKKKTKTGKTKKTTKTTKKTTTTKKTKTKKTKTKGKKKGKGKKKTBKGG 583

Qy      96 GCAAGCAGCTGCGACACATTCTGGGGGGTGATTACGGAGTCCCAACATTTTCTAACGGTGG 155
           :  |  | :| |      :  :|||||   ::  :  :|       ::|: :     :::
Db     584 TBGGGGKGKBGGGGGGGGGKGGKGGGGGGTTKKKTKTTKTKKTKTTTKKTKTKKTKTTKKK 643

Qy     156 TATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCGTGCTTACGG 215
           :   | : : |   :: :     :    |: :::|  |  |     :  : |||::| |:|
Db     644 KKGTTKYTKKGBKTKKTBTCCKKKTTTTBKKBKTTTTKTTKTKTKYKTKTGTGBBTYCSG 703

Qy     216 GTGCGGTGGGTCTGGGGGTAGCGTTAGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGC 275
           :|:::::::::: ::|||:  | ||: :| ::   :|::: : ::   :: ||:| ||
Db     704 KTKBBKKKKKKGKBGGGKGKGKGTBKBTKKKKYTBTBBKYBTBBYKTKYBTTKTTTTBTT 763

Qy     276 TATTGTTTCTGTGTT 290
           : ||:::| |:: ::
Db     764 BCTTBYBTTTKYTYB 778


RESULT 4
BJ032222
LOCUS       BJ032222                598 bp    mRNA    linear   EST 05-DEC-2001
DEFINITION  BJ032222 NIBB Mochii normalized Xenopus neurula library Xenopus
            laevis cDNA clone XL016a10 5', mRNA sequence.
ACCESSION   BJ032222
VERSION     BJ032222.1  GI:17378218
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 598)
  AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
            ,Y.
  TITLE     Expressed genes in X. laevis embryo
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .598
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="XL016a10"
                     /clone_lib="NIBB Mochii normalized Xenopus neurula
                     library"
                     /tissue_type="whole embryo"
                     /dev_stage="stage 15"
                     /note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
                     were oligo-dT primed and directionally cloned. Staging
                     according to  Nieuwkoop and Faber. Library is subtracted
                     and was constructed by N. Garrett  and A.M. Zorn,
                     (Wellcome/CRC Institute). "
BASE COUNT      138 a    130 c    165 g    165 t
ORIGIN

  Query Match           11.0%;  Score 39.2;  DB 13;  Length 598;
  Best Local Similarity  58.6%;  Pred. No. 8.6;
  Matches   68;  Conservative    0;  Mismatches   48;  Indels    0;  Gaps    0;

Qy      36 TGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCA 95
           ||||||| | |||||||| ||| ||||| |||   |||||||| |||| | ||
Db     249 TGTGCCTGCAGCTGTTGGTTTTTCTGTTGGGGCAGGCTTTGTATACTTGGCCGACCAGCT 308

Qy      96 GCAAGCAGCTGCGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACG 151
           | | |||||  |       ||| || |  |||  ||  | ||| |  | | |
Db     309 ACTACCAGCTCTGGGCTTCACTGAGGATCCCTACTCCGTTGCCACACTGAATCAGG 364


RESULT 5
CNS00FOO
LOCUS       CNS00FOO                1101 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence T7 end of BAC:
            BACR31F03 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL070854
VERSION     AL070854.1  GI:4950896
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACR31F03"
                     /clone_lib="RPCI-98"
                     /note="end : T7"
BASE COUNT      167 a     84 c     86 g    155 t    609 others
ORIGIN

  Query Match           10.8%;  Score 38.4;  DB 17;  Length 1101;
  Best Local Similarity  14.6%;  Pred. No. 15;
  Matches   43;  Conservative  129;  Mismatches  122;  Indels    0;  Gaps    0;

Qy      22 CGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTG 81
           :| ::::  ::  ::|::    |:::   ::::  |: :   ||:  :   :||:| :|:
Db     804 MGGMMMMGAMMTGVKGKKGTGGTKVBGTGYKKKNGTKNKMNMGTKWTDMNTKTTKTTKTK 863

Qy      82 CTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGGGGTGATTACGGAGTCCCAACA 141
             ::   |:  |:       :  : :: :|   :::  : : :: | :  ::|:    |: |
Db     864 TKKTTATWTKTKTGMKKKDKKKKKKTKAKKKMKKKAKAKAKKMAGDGAKDAKATAAAWAA 923
```

```
Qy     142 TTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAAT 201
Db     924 KAKAKAADAAAAWAAADAAAKAADKGTRWDGAKAKRTKKDAKKARAKADAKADDRAKKDK 983

Qy     202 CATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAGTAGTCGAGTTGGGCAGCAA 261
Db     984 WKKKKKKKKKKRDKKGKKDKKGKDKDKKKKGAKAKAKKKATKKKKDKAKKDKAKDKAKKR 1043

Qy     262 CTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCACTATTACAACGATTA 315
Db    1044 KDKKAKWKTKTWDATDADTDKKAKAKDKTWKDKKWKTATAKATDATAKAKWATA 1097
```

RESULT 6
CNS006U0
LOCUS CNS006U0 884 bp DNA linear GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC # BACR14N21 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL065923
VERSION AL065923.1 GI:4944891
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 884)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES Location/Qualifiers
source 1. .884
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR14N21"
/clone_lib="RPCI-98"
/note="end : T7"
BASE COUNT 230 a 62 c 139 g 124 t 329 others
ORIGIN

Query Match 10.6%; Score 37.8; DB 17; Length 884;
Best Local Similarity 11.0%; Pred. No. 20;
Matches 25; Conservative 113; Mismatches 90; Indels 0; Gaps 0;

```
Qy      45 TGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGCAGC 104
Db     654 KGGGGGGKKKGGTKTGGKDGTGGKKKTTKKKKKKKKGTKGKGKGKTKKKSGTKKKKGGTG 713

Qy     105 TGCGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAG 164
Db     714 KGTGTKGTGTKKKKGGKGKGGKKKGGKGKKGKKKKKGKKKKKKKKKGKKKKKTGKKTKGK 773

Qy     165 AGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGG 224
```

```
Db     774 KGGTKKKGKKKGKKTKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKTKKKKKKKK 833

Qy     225 GTCTGGGGGTAGCGTTAGTAGTCGAGTTGGGCAGCAACTTATTGTGTT 272
Db     834 KKCKKKKKKKKKKKKKKKKKKKKKKKNKBKKDAAMMMMMADKKKBKKK 881
```

RESULT 7
CNS0100X
LOCUS CNS0100X 1101 bp DNA linear GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence SP6 end of BAC BACN03G04 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL098379
VERSION AL098379.1 GI:5609990
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBeloBAC11.
FEATURES Location/Qualifiers
source 1. .1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACN03G04"
/clone_lib="DrosBAC"
/plasmid="pBeloBAC11"
/note="end : SP6"
BASE COUNT 195 a 108 c 131 g 161 t 506 others
ORIGIN

Query Match 10.5%; Score 37.6; DB 17; Length 1101;
Best Local Similarity 17.4%; Pred. No. 24;
Matches 55; Conservative 132; Mismatches 129; Indels 0; Gaps 0;

```
Qy      42 TATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCAAGC 101
Db     506 TWTDVWADNAGWWTDKDKTKAKKKRTKTDTKKTKTRTGTKKKTDRGMRKTDTWGWGDWRA 565

Qy     102 AGCTGCGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATA 161
Db     566 DRTDGARADRAANAKDKDADGKADTDATKTDGWGTTTKADTTDDTDADWWDWKDWDWKDA 625

Qy     162 TAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGG 221
Db     626 DTKAGRKRGWGWGKKGGWKKTTKTKKKKTDKTKTGTDTKWGWKDKWTDKKRDGKRGDGDR 685

Qy     222 TGGGTCTGGGGGTAGCGTTAGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGT 281
Db     686 KDKGTGGDGDSKAKGDRAWAWRAKATRAAAADATAAATAAKATAWATKTTTKKTTTTTTT 745

Qy     282 TTCTGTGTTAATAGTGTCACTATTACAACGATTAAGGTCTCCACCAGAACACATTTGTAA 341
Db     746 KTTTTTTTTTTTKKTTTARKADRDDDDKAKRDGDKDGDWDDADKDWKAGRDRDDDTRKWG 805

Qy     342 TGGTGCTTGTGGTTAA 357
Db     806 WKDRAGKKKAKKKDH 821
```

```
RESULT 8
BB671302/c
LOCUS       BB671302                401 bp    mRNA    linear   EST 05-OCT-2001
DEFINITION  BB671302 RIKEN full-length enriched, 10 days embryo Mus musculus
            cDNA clone 3426401D14 3', mRNA sequence.
ACCESSION   BB671302
VERSION     BB671302.1  GI:15970523
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 401)
  AUTHORS   Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
            Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii
            ,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
            Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.
            , Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
            Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
            ,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.
            , Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
            2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .401
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="3426401D14"
                     /clone_lib="RIKEN full-length enriched, 10 days embryo"
                     /sex="mixed"
                     /dev_stage="10 days embryo"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. Second strand cDNA was prepared with the
                     primer adapter of sequence [5'
                     GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA
                     was cloned into the XhoI and BamHI sites. Vector: a
                     modified pBluescript KS(+) after bulk excision from Lambda
                     FLC I"
BASE COUNT      160 a     45 c     65 g    131 t
ORIGIN

  Query Match             10.5%;  Score 37.4;  DB 10;  Length 401;
  Best Local Similarity   50.9%;  Pred. No. 24;
  Matches   89;  Conservative    0;  Mismatches   86;  Indels    0;  Gaps    0;

Qy      26 CCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGG 85
           ||| | | | | || || | ||  |   |   | | |     ||||  ||||  |||
Db     230 CCATTGACATTTTGTCTTTCGTCTTCTTTCGGTTTTCTACAATTTTCAACGTATATCTGT 171

Qy      86 CGTTCATGCAGCAAGCAGCTGCGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTT 145
             || || | ||      |||| | |||||    |  || || |  | ||  |  ||| |
Db     170 ATTTAATACTGCTTTATTCTGCTAGACATTAACTGACTGCTTTCTAAATCTAAGGATTAT 111

Qy     146 CTAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAA 200
            ||    | ||||| ||| | |    | |   ||     ||||| | |  | |||
Db     110 ATAGTTTTAGTATAAATATATATCACATATTTTCGTAAAATTTTGACAAAACCAA 56


RESULT 9
BF417205
LOCUS       BF417205                474 bp    mRNA    linear   EST 28-NOV-2000
DEFINITION  UI-R-CN0-blc-c-11-0-UI.s1 UI-R-CN0 Rattus norvegicus cDNA clone
            UI-R-CN0-blc-c-11-0-UI 3', mRNA sequence.
ACCESSION   BF417205
VERSION     BF417205.1  GI:11405194
KEYWORDS    EST.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 474)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
COMMENT     Contact: Soares, MB
            Program for Rat Gene Discovery and Mapping
            University of Iowa
            451 Eckstein Medical Research Building Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: msoares@blue.weeg.uiowa.edu
            Oligo-dT track not found, Not I site shown in beginning of sequence
            is likely internal to the message. cDNA Library Preparation: M.B.
            Soares Lab Clone distribution: clones will be available through
            Research Genetics (www.resgen.com)
            Seq primer: M13 Forward
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .474
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /clone="UI-R-CN0-blc-c-11-0-UI"
                     /clone_lib="UI-R-CN0"
                     /dev_stage="ADULT"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-CN0
                     library is a normalized library constructed from the
                     following tissues: brown adipose, penis, salivary glands,
```

bladder, fundus, cervix, seminal vesicles. For a detailed description of the library from which this clone was derived, please visit our web site at ratest.eng.uiowa.edu. The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)
TAG_SEQ=None found"
BASE COUNT 118 a 134 c 98 g 124 t
ORIGIN

Query Match 10.4%; Score 37.2; DB 12; Length 474;
Best Local Similarity 55.4%; Pred. No. 27;
Matches 72; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

```
Qy      39 GCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAGCA 98
           |||| || ||||| || | | |||| |   | | |   || || || |     | ||||
Db      11 GCCTTTTTTTGTTTGTTTGTTTGTTTTTAATATATAGCTACTGATGCCTCAGGTCCAGCC 70

Qy      99 AGCAGCTGCGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTAT 158
           | |||  ||  | ||  |   |||| | |    |||| | ||  |||   |  |   | |
Db      71 ACCAGGAGCATCTCACCCGACGGGTAAATCTCTAGTCACGACCCTTTTAGAAAGACATGT 130

Qy     159 ATATAGAGAC 168
           | ||| | ||
Db     131 AAATATATAC 140
```

RESULT 10
BQ215809
LOCUS BQ215809 873 bp mRNA linear EST 02-MAY-2002
DEFINITION AGENCOURT_7574890 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6059837 5', mRNA sequence.
ACCESSION BQ215809
VERSION BQ215809.1 GI:20397219
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 873)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13327 row: c column: 06
High quality sequence stop: 645.
FEATURES Location/Qualifiers
source 1. .873
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6059837"
/clone_lib="NIH_MGC_68"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.8 kb. Library constructed by Life Technologies. "
BASE COUNT 203 a 224 c 162 g 284 t
ORIGIN

Query Match 10.3%; Score 36.6; DB 14; Length 873;
Best Local Similarity 49.7%; Pred. No. 41;
Matches 93; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

```
Qy     123 TGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGC 182
           |||   |   |  ||||||    | |    ||  |  |  | || | |  | | ||  ||
Db      33 TGACCCCCTGGCACCAACACCCGCCAGTACTGTGACTTCCAAAGCCAGAGCCACATGTGC 92

Qy     183 TGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAG 242
           | ||   | | | |  || ||    ||    |   || ||||  ||||||||     |
Db      93 TCATCAAACTTGCATTAAGCAGTTGGCGGGAGATGGCTGTGGAGCTGGGGGTTTAAGTGA 152

Qy     243 TAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCACT 302
           | ||    ||  ||  |  ||| |   | ||    |  | ||| | |||| |||||   |
Db     153 TGGTTCTCTTTTGCTCCCTCTTTTGAGGGTAAAGCTACTGTCTTTCTTAAGAGTGTATTT 212

Qy     303 ATTACAA 309
           || |||
Db     213 ATGCCAA 219
```

RESULT 11
AA134045
LOCUS AA134045 486 bp mRNA linear EST 27-NOV-1996
DEFINITION zo17d11.r1 Stratagene colon (#937204) Homo sapiens cDNA clone IMAGE:587157 5', mRNA sequence.
ACCESSION AA134045
VERSION AA134045.1 GI:1691275
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 486)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 318.
FEATURES Location/Qualifiers
source 1. .486
/organism="Homo sapiens"
/db_xref="GDB:4619459"
/db_xref="taxon:9606"
/clone="IMAGE:587157"
/clone_lib="Stratagene colon (#937204)"
/tissue_type="tumor"
/cell_line="T84 carcinoma cell line"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: colon; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. T-84 colonic epithelial cell line. Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTTT 3'"
BASE COUNT 136 a 119 c 89 g 138 t 4 others
ORIGIN

Query Match 10.2%; Score 36.4; DB 9; Length 486;
Best Local Similarity 50.6%; Pred. No. 44;
Matches 88; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

```
Qy     136 CCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGT 195
```

```
          ||||||    | |    ||  |  |  | || | |  | | ||  ||| ||   | | |
Db    119 CCAACACCCGCCAGTACTGTGACTTCCAAAGCCAGAGCCACATGTGCTCATCAAACTTGC 178

Qy    196 AACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAGTAGTCGAGTTGGG 255
          |  || ||    ||    |  || ||||  ||||||||     |  | ||    ||  |
Db    179 ATTAAGCAGTTGGCGGGAGATGGCTGTGGAGCTGGGGGTTTAAGTGATGGTTCTCTTTTG 238

Qy    256 CAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGTGTCACTATTACAA 309
          |  |  ||| |   | ||    |  | ||| | |||| |||||   |||  |||
Db    239 CTCCCTCTTTTGAGGGTAAAGCTACTGTCTTTCTTAAGAGTGTATTTATGCCAA 292
```

RESULT 12
CNS0073W

```
LOCUS       CNS0073W                922 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR14D09 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL066784
VERSION     AL066784.1  GI:4945247
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 922)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .922
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACR14D09"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
BASE COUNT      223 a     95 c    109 g    221 t    274 others
ORIGIN

  Query Match           10.1%;  Score 36.2;  DB 17;  Length 922;
  Best Local Similarity   16.8%;  Pred. No. 53;
  Matches   44;  Conservative  102;  Mismatches  116;  Indels    0;  Gaps    0;

Qy      36 TGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCA 95
           :|::  | : ::|||||| :::|:: :::::  ::: ::||:  :: |:: :|:  :|
Db     640 KGKKKGTGKGKKTGTTGGGKKKTKKTKKKKKTGKKKKKKTGKKGKKGTKKTKTKTTKGGK 699

Qy      96 GCAAGCAGCTGCGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACGGTGG 155
           |      : |  :      |: ::::: |: ::   : ::          : | :   |:::|
Db     700 GGKKTKKKKTTKKTTKTTKTKKKKKTTKTKKTGTKKKKTTKTTKKGKGTKKGKKGKKGG 759

Qy     156 TATATATAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCGTGCTTACGG 215
           : : | | : |  |:     : :  : ::  :::    ||      |  | :|: |:  ::
Db     760 KTKGTGTKKKGTGGKGKKGKKGKKKKKKKKGKKKKTGKGTGTKKKTKGTGKTKKTKKKKK 819

Qy     216 GTGCGGTGGGTCTGGGGGTAGCGTTAGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGC 275
           ::: : :::|| :|||: ::   : : ::  : : :||::   |    :: ::|:|:  :
Db     820 KKKTKTKKKGTKKGGKTKKGTKKGKKKKGTKKKNKTTKKTGTGTKKKKKKKKGKGKGKKK 879

Qy     276 TATTGTTTCTGTGTTAATAGTG 297
           :  ::: : :|:| :  : |::
Db     880 KKGKKKGKKKGKGGKGKKKGKK 901
```

RESULT 13
BB224881

```
LOCUS       BB224881                328 bp    mRNA    linear   EST 01-JUL-2000
DEFINITION  BB224881 RIKEN full-length enriched, adult male aorta and vein Mus
            musculus cDNA clone A530088K07 3' similar to D43951 Homo sapiens
            mRNA for KIAA0099 protein, mRNA sequence.
ACCESSION   BB224881
VERSION     BB224881.1  GI:8893493
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 328)
  AUTHORS   Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci
            ,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
            Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
            Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
            Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M.
            , Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.
            , Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata
            ,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y.
            , Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya
            ,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I.
            , Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino
            ,M., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Konno,H., et al.)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki
            ,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Thermostabilization and thermoactivation of thermolabile enzymes by
            trehalose and its application for the synthesis of full length
            cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
             Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
            Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
            ,Y. and Hayashizaki,Y.
             Automated filtration-based high-throughput plasmid preparation
            system. Genome Res. 9 (5), 463-470 (1999)
             Carninci,P. and Hayashizaki,Y.
             High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
            19-44 (1999)
             Please visit our web site (http://genome.rtc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .328
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="A530088K07"
                     /clone_lib="RIKEN full-length enriched, adult male aorta
                     and vein"
                     /sex="male"
                     /tissue_type="aorta and vein"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
```

prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT 83 a 70 c 50 g 125 t
ORIGIN

Query Match 10.1%; Score 36; DB 10; Length 328;
Best Local Similarity 51.9%; Pred. No. 53;
Matches 81; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

```
Qy    177 ATCAGCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAG 236
          ||||  |  ||| |  | |||    | || || | || | ||    ||  || |
Db     82 ATCATTTTTTTTAATGACTAATTCCCCTCCTGTTCACTGCTGTTTAGGTACTTGTACATA 141

Qy    237 CGTTAGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGTTTCTGTGTTAATAGT 296
          ||||  |||  | |||||  || ||||  | ||||   |||||| ||| | ||     |
Db    142 GGTTATCAGTGTACTTGGGTGGCCACTTGCTTTGTTCTATATTGTATCTTTATTCCACAT 201

Qy    297 GTCACTATTACAACGATTAAGGTCTCCACCAGAACA 332
          ||    | |   | | |    |  |  || | || |
Db    202 GTGGACACTTTCAGGGTGGTTGGTTTAACAAAAAAA 237
```

RESULT 14
AQ927406/c
LOCUS AQ927406 498 bp DNA linear GSS 21-DEC-1999
DEFINITION RPCI-23-261L19.TV RPCI-23 Mus musculus genomic clone RPCI-23-261L19 , DNA sequence.
ACCESSION AQ927406
VERSION AQ927406.1 GI:6616409
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 498)
AUTHORS Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret ,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P. and Fraser,C.M.
TITLE Mouse BAC End Sequences from Library RPCI-23
JOURNAL Unpublished (1999)
COMMENT Other_GSSs: RPCI-23-261L19.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 261 row: L column: 19
Seq primer: T7
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .498
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-261L19"
/clone_lib="RPCI-23"
/sex="Female"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 190 a 196 c 36 g 75 t 1 others
ORIGIN

Query Match 10.0%; Score 35.8; DB 17; Length 498;
Best Local Similarity 49.4%; Pred. No. 62;
Matches 121; Conservative 0; Mismatches 122; Indels 2; Gaps 1;

```
Qy     44 TTGTTGTTGGTGTTTGTGTTGTGGCTTTCTTTGTATTGCTGGCGTTCATGCAG--CAAGC 101
          ||||||||||  ||  ||||||| | | |  | ||  || ||| | |  || ||     |
Db    409 TTGTTGTTGGGTTTGTTGTTGTTGTGGTGGTGGTGGTGGTGGTGGAGGTGGTGG 350

Qy    102 AGCTGCGACACATTCTGGGGGTGATTACGGAGTCCCAACATTTTCTAACGGTGGTATATA 161
           | ||      ||  ||  |||| | | || |          |  |   ||||||
Db    349 TGGTGGTGGTGATGGTGTTGGTGGTAATGGTGGTGGTGATGATGGTGGTGGTGGTGG 290

Qy    162 TAGAGACGGTACAAGATCAGCTGATTTTAATAGTAACAATCATCGTGCTTACGGGTGCGG 221
          | | || |||    |    | || |  |  |  | |   |  | ||| | | ||  | ||
Db    289 TGGTGATGGTGGTGGTGATGATGGTGGTGGTGATGATGGTGGTGGTGATGATGGTGGTGG 230

Qy    222 TGGGTCTGGGGGTAGCGTTAGTAGTCGAGTTGGGCAGCAACTTATTGTGTTAGCTATTGT 281
          |||   ||| ||| | | |  |||| | | |||     ||   | ||     | |  ||
Db    229 TGGTGATGGTGGTGGTGGTGATAGTGGTGGTGGTGGTGAAGAAAGTGGTGGTGGTGATGG 170

Qy    282 TTCTG 286
          |  ||
Db    169 TGGTG 165
```

RESULT 15
CNS00CNG/c
LOCUS CNS00CNG 939 bp DNA linear GSS 04-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR26H16 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL059400
VERSION AL059400.1 GI:4946964
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 939)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's

```
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .939
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACR26H16"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
BASE COUNT       71 a    349 c    104 g    180 t    235 others
ORIGIN

  Query Match            10.0%;  Score 35.8;  DB 17;  Length 939;
  Best Local Similarity  11.3%;  Pred. No. 67;
  Matches   26;  Conservative  115;  Mismatches   90;  Indels    0;  Gaps    0;

Qy      12 AATAACCGCTCGACCCAATAAGAATGTGCCTATTGTTGTTGGTGTTTGTGTTGTGGCTTT 71
           || ||  ||  | | |   |   |    |  |:  :::::::::::::::::: :::
Db     603 AAGAAGGGCAGGGCACGGGAGAGAGAGACGGAKAAKKKKKKGKKKKKKDDKKKKKKKKK 544

Qy      72 CTTTGTATTGCTGGCGTTCATGCAGCAAGCAGCTGCGACACATTCTGGGGGTGATTACGG 131
            :::::|::: ::: :::  ::  :  :  :  : :     :: : :::::: ::: ::
Db     543 KDKKKKAKKKKKKKKKKAKKKAKKKKKKKKDDADKKKKKDKKKKKAKDDKKKKKKDKKD 484

Qy     132 AGTCCCAACATTTTCTAACGGTGGTATATATAGAGACGGTACAAGATCAGCTGATTTTAA 191
           |::      |:::: :   :::::: : : : :|:   ::    :   |: ::|:|:: |
Db     483 AKKKKKKKKAKKKKKKKKAKKKKKKKKKKKKKAKKAAKKKKKKKKKAKAKGKKAKTDKKA 424

Qy     192 TAGTAACAATCATCGTGCTTACGGGTGCGGTGGGTCTGGGGGTAGCGTTAG 242
           :: :  |:  : ::: ::  ::::: ::::::: ::|| :: : : : |
Db     423 AKKKKDKKAKKKKKKKKKKDKAKKKKKKKKKKKKKKKKKGGAKDGKKKGKGG 373


Search completed: March 12, 2003, 10:11:21
Job time : 1196.5 secs
```